



# 基于互联网群体智能的拼图问题求解

申博<sup>1,2</sup>, 张伟<sup>1,2\*</sup>, 赵海燕<sup>1,2</sup>, 金芝<sup>1,2\*</sup>, 吴艳红<sup>3</sup>

1. 高可信软件技术(北京大学)教育部重点实验室, 北京 100871

2. 北京大学信息科学技术学院软件研究所, 北京 100871

3. 北京大学心理与认知科学学院, 北京 100871

\* 通信作者. E-mail: zhangw.sei@pku.edu.cn, zhijin@pku.edu.cn

收稿日期: 2019-02-25; 修回日期: 2019-07-15; 接受日期: 2019-09-24; 网络出版日期: 2021-02-01

国家重点基础研究发展计划(973计划)(批准号: 2015CB352201)和国家自然科学基金(批准号: 61690200, 61432020)资助项目

**摘要** 拼图游戏是一类复杂问题的典型代表. 这类问题的复杂性体现在其无法通过自上而下、集中控制的方式进行有效求解. 通常情况下, 拼图游戏由一个玩家独立完成, 或由几个玩家围坐在一起协同完成. 本文提出一种基于互联网人类群体智能的拼图问题求解方法. 该方法使用一种被称为“探索-融合-反馈”的回路用于支持人类玩家群体在互联网环境中进行并行、协同的拼图问题求解. 其中, 探索活动由玩家实施, 融合与反馈活动由拼图环境自动实施. 在探索活动中, 每个玩家独立进行拼图问题的求解, 不与其他玩家发生直接的交互. 在任何时刻, 一个玩家探索活动的结果是一个关于当前拼图问题的片段解, 以及一组被该玩家判断为错误的图块邻接关系. 在融合活动中, 拼图环境实时地将所有玩家的当前探索结果融合在一起, 形成一个不断演化的群体观点图. 在反馈活动中, 对于每一个玩家, 拼图环境根据该玩家的当前探索结果和当前的群体观点图, 向该玩家推荐特定的信息, 以加速其拼图进程. 我们实现了一个支持上述方法的多人在线拼图环境, 并在 2~10 人规模的玩家群体中进行了初步的实验. 实验结果表明: (1) 在这一多人拼图环境中, 拼图求解时间大致正比于玩家数量的倒数, 且以玩家群体中的最强个体为基准点, 群体拼图效率能够实现 31.36%~64.57% 的提升; (2) 在该环境中, 最快完成拼图的玩家获得的反馈信息平均具有 86.34% 的准确率, 且随着群体规模的增加, 反馈信息在最快玩家拼图结果中的平均占比逐渐从 20% 增加到 45% 左右; (3) 相比于面对面协同的群体拼图求解方式, 该环境展示出更好的群体规模可扩展性, 且拼图结果总是具有 100% 的正确率, 而拼图问题自动求解算法平均只具有 52% 的正确率. 我们希望本文工作能够为探索群体智能在互联网环境下的更广泛应用提供一些有用的信息或观点.

**关键词** 人类群体智能, 互联网, 复杂问题求解, 拼图问题, 信息融合与反馈

**引用格式:** 申博, 张伟, 赵海燕, 等. 基于互联网群体智能的拼图问题求解. 中国科学: 信息科学, 2021, 51: 206-230, doi: 10.1360/SSI-2019-0150

Shen B, Zhang W, Zhao H Y, et al. Solving pictorial jigsaw puzzles via Internet-based collective intelligence (in Chinese). Sci Sin Inform, 2021, 51: 206-230, doi: 10.1360/SSI-2019-0150

## 1 引言

对人类而言, 拼图是一款经典的休闲游戏. 在这个游戏中, 一幅图片被事先分割成一组乱序的图块. 玩家需要通过拼接这些图块尽快恢复原图<sup>1)</sup>. 在游戏的表象之下, 拼图具有深刻的象征意义, 是一类复杂问题的典型代表. 这类问题的复杂性体现在其无法通过自上而下、集中控制的方式进行求解, 即: 无法事先确定其严格求解过程, 进而也无法在求解过程的控制下对问题进行有序地求解. 这类问题只能采用不断探索、不断组合与不断优化的方式进行求解, 即: 不断探索问题相关的信息片段; 尝试对这些信息片段进行各种可能的组合, 遴选其中具有潜在价值的信息片段组合体; 对这些信息片段组合体进行持续的改进、优化, 以及进一步的组合, 直到形成一个具有足够高质量的解决方案. 本文将这类复杂问题称为“拼图类问题”.

拼图类问题在各领域中广泛存在. 例如, 在地质学领域中, 有学者基于地质数据, 尝试采用拼图的方式还原远古时期的超级大陆<sup>[1]</sup>. 在医学领域中, 有学者认为, 如果能把人类关于大肠癌的各种知识片段拼接在一起, 就有可能发现大肠癌的形成机理<sup>[2]</sup>. 在生物学领域中, 有学者将核糖体 RNA 的结构理解为一种三维拼图问题<sup>[3]</sup>. 现实世界中也充满了各种无法采用自上而下、集中控制的方式进行求解的拼图类问题<sup>[4]</sup>.

我们选择拼图问题作为研究关注点, 主要基于两点考虑: (1) 从一种新的视角探索如何高效求解拼图问题; (2) 在此基础上, 为解决更一般的“拼图类问题”奠定必要的研究基础. 短期而言, 我们想探索一种基于互联网人类群体智能的拼图问题求解方法. 中期而言, 我们想把从群智拼图研究中探索出的行之有效的方法和技术应用到具有复杂逻辑结构的信息制品的群智化构造中. 例如, 采用群体智能的方式构造特定领域内规模巨大的知识图谱、采用群体智能的方式进行复杂软件系统的开发<sup>[5]</sup>等. 长期而言, 我们想证明, 群体智能及其所蕴含的组织机理, 为人类文明在信息社会中的进一步发展指出了一种方向, 使得在机器和算法智能蓬勃发展的情况下, 人类这一生物物种仍然具有存在的价值和意义.

从算法的角度看, 拼图问题是一种非良好定义的问题 (ill-defined problem), 因为拼图问题的定义中并不包含如何判断一个候选方案正确与否、或质量高低的量化准则. 缺少对候选方案正确性/质量的量化判断准则, 使得仅靠算法很难在一般意义上实现对拼图问题的有效求解. 虽然人类能够容易地对任何一个候选拼图方案的正确性/质量进行主观判断, 但这并不意味着拼图问题对人类而言是一种简单的问题. 在我们的实验中, 人类个体求解一个包含 100 个图块片段 (10 行  $\times$  10 列) 的拼图问题, 平均也需要花费大约 24 分钟的时间. 其主要困难在于该问题具有一个大规模的候选方案空间. 虽然人类可以利用其对图片内容的感知能力对这一候选方案空间进行一定程度的缩减, 但缩减后的空间, 依靠单一人类个体的信息处理能力仍然无法被有效地检索.

通常情况下, 拼图问题由一个玩家独立求解, 或由几个玩家围坐在一起协同求解. 本文主要关注一组在地理上分布的人类玩家如何在互联网的支持下实现对拼图问题的高效求解. 我们在环境激发效应 (stigmergy) 概念的启发下, 提出了一种基于互联网人类群体智能的拼图问题求解方法.

环境激发效应于 1959 年由法国动物学家 Pierre-Paul Grassé 提出<sup>[6~9]</sup>, 表示昆虫个体之间以环境为媒介的间接交互机制, 用于解释在社会性昆虫群体中通常会观察到的一种矛盾现象: 在个体层次上, 每一个昆虫看起来都在独立地进行自己的活动, 个体之间似乎没有进行任何交互; 但在群体层次上, 所有个体的行为综合在一起, 却体现出高度的协同性. 这种现象被研究者称为社会性昆虫群体中的群体智能 (swarm/collective intelligence, CI) 现象. 在环境激发效应的视角中, 昆虫群体的群体智能现象通

1) 在本文中, 假设在拼图游戏开始前, 玩家看不到原图.

过如下方式涌现形成: 昆虫个体在环境中留下自己的踪迹或对环境进行某种形式的修改; 然后, 这些踪迹或修改被群体中的个体所感知, 并刺激个体在环境中留下新的踪迹或对环境进行进一步的修改; 不同个体的行为通过这种方式相互协同, 形成一个正反馈回路, 进而导致智能化的自组织群体行为的涌现.

环境激发效应揭示了群体智能中的一个关键构成成分: 环境. 正是在环境的支持下, 才有可能形成个体之间以环境为媒介的大规模间接交互行为. 基于环境激发效应的群体智能具有一个重要的特点: 良好的群体规模可扩展性. 在环境容量允许的情况下, 任何规模的群体中的个体之间都可以基于环境激发效应展现出有效的大规模群体协同行为.

以环境激发效应为指导, 设计基于环境激发效应的人工群体智能系统 (artificial CI system) 的一个关键点是构造一种适应待求解问题特点的虚拟环境. 我们认为, 环境激发效应中的环境应该承担两种职责. 第 1 种是融合: 环境能够将群体中不同个体提供的信息片段融合在一起, 形成一个良构的群体层制品. 第 2 种是反馈: 环境能够基于当前的群体层制品为不同的个体提供有针对性的反馈信息, 以提高其工作效率. 在那些发生在物理空间中的基于环境激发效应的自然群体智能系统 (natural CI system) 中, 融合通常由物理规律承担. 例如, 在物理规律的作用下, 蚁群中的不同个体在同一位置释放出的信息素 (一种化学物质) 被融合在一起, 导致该位置上信息素浓度的增加; 同样, 在物理规律的作用下, 一个位置上的信息素会因为挥发效应的存在而使得其浓度逐渐降低. 在这些自然群体智能系统中, 环境会根据不同个体在物理空间中的位置为其反馈不同的信息 (即: 个体在不同的物理位置上会接受到不同的环境信息). 但是, 在信息空间中, 物理规律对信息的处理不再具有直接的效果; 同时, 物理位置也不再具有意义. 因此, 为了设计信息空间中基于环境激发效应的人工群体智能系统, 需要构造出一种适应信息空间和待求解问题特点的信息融合与反馈机制.

本文方法的核心是一个持续运行的回路, 称为“探索 - 融合 - 反馈”回路 (见图 1). 该回路包含了 3 个并行的活动: 探索、融合、反馈. 其中, 前者由人类玩家实施, 后两者由拼图环境自动实施. 在探索活动中, 每个玩家独立进行拼图问题的求解, 不与其他玩家发生直接的交互. 在任何时刻, 一个玩家的探索活动的输出是一个关于当前拼图问题的片段解, 以及一组被该玩家判断为错误的图块邻接关系. 在融合活动中, 拼图环境实时地将所有玩家的当前探索结果融合在一起, 形成一个不断演化的群体观点图. 在反馈活动中, 对于每一个玩家, 拼图环境根据该玩家的当前探索结果和当前的群体观点图, 向该玩家推荐特定的信息, 以加快其拼图进程. 每一个玩家自主决定接受或拒绝拼图环境提供的反馈信息. 当任何一个玩家发现了正确的拼图答案, 则表明当前拼图问题已经得到了求解. 此时, 可以选择终止“探索 - 融合 - 反馈”回路.

“探索 - 融合 - 反馈”回路是为了实现人类群体在信息空间中的群体智能而对环境激发效应这一概念的具体化. 这种具体化体现在两个方面. 首先, 该回路中的融合不再由物理规律所承担, 而是由运行在计算机上的人工算法所承担; 这种改变对于将环境激发效应应用于不同的问题求解场景带来了充分的灵活性. 第二, 该回路中的反馈不再根据个体在物理空间中的位置, 而是基于当前的群体层制品和个体的当前探索结果进行个性化的信息反馈.

我们实现了一个支持“探索 - 融合 - 反馈”回路的多玩家在线拼图环境<sup>2)</sup>. 基于该环境, 我们在 2~10 人规模的玩家群体中进行了一系列的受控实验. 实验主要关注 3 个具体的研究问题: (1) 以拼图问题规模和玩家群体规模为两个独立的自变量、以群体性能为因变量, 建立自变量与因变量之间的经验型因果关系; (2) 观察在不同的拼图问题规模和玩家群体规模的组合下, 群体中 fastest 完成拼图的玩家

2) 该环境的网址为 <http://www.pintu.fun>, 目前仅支持 Chrome 浏览器.

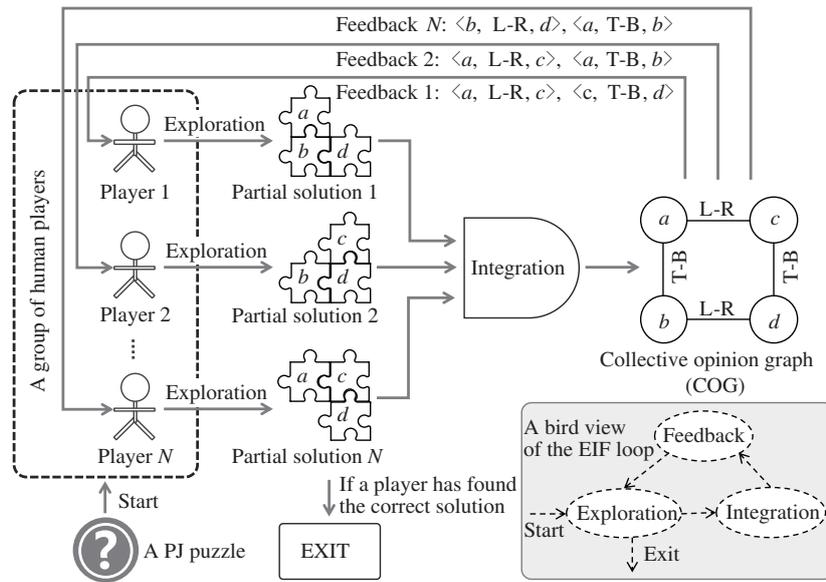


图 1 基于“探索 - 融合 - 反馈”的群体协同式拼图问题求解

Figure 1 Solving the PJ puzzle collaboratively based on the explore-integration-feedback (EIF) loop

接受到的反馈信息的准确率, 以及反馈信息在此类玩家拼图结果中所占的比例, 验证本文提出的融合与反馈算法的有效性; (3) 将本文方法与面对面群体拼图方法, 以及拼图问题自动求解算法进行对比, 考察本文方法的优缺点. 我们期望通过对上述 3 个问题的实验性研究, 证明本文提出的基于“探索 - 融合 - 反馈”回路的群体拼图方法是一种有效的群智化拼图方法. 实验结果表明:

(1) 在这一多人拼图环境中, 拼图求解时间大致正比于玩家数量的倒数, 且以玩家群体中的最强个体为基准点, 群体拼图效率能够实现 31.36%~64.57% 的提升;

(2) 在该环境中, 最快完成拼图的玩家获得的反馈信息平均具有 86.34% 的准确率, 且随着群体规模的增加, 反馈信息在最快玩家拼图结果中的平均占比逐渐从 20% 增加到 45% 左右;

(3) 相比于面对面协同的群体拼图求解方式, 该环境展示出更好的群体规模可扩展性, 且拼图结果总是具有 100% 的正确率, 而拼图问题自动求解算法平均只具有 52% 的正确率.

本文的主要贡献包含如下 4 点:

(1) 一个针对拼图问题、基于环境激发效应的人类群体智能概念模型; 称为“探索 - 融合 - 反馈”回路.

(2) 一个能够将拼图求解过程中所有个体的探索结果实时融合在一起的结构化制品模型; 称为群体观点图.

(3) 一个支持玩家群体通过协同和去中心化的方式求解拼图问题的多人在线拼图环境.

(4) 一个关于群体规模、问题规模这两个因素与群体拼图性能之间因果关系的经验性量化模型.

本文余下部分的安排如下. 第 2 节简要介绍本文方法的重要相关工作. 第 3 节论述一种在信息空间中支持多玩家协同并行求解拼图问题的方法. 第 4 节通过一组受控实验对本文方法的可行性和效果进行评估. 第 5 节对本文方法相关的若干基础性问题进行简要讨论. 第 6 节对本文工作进行总结, 并对下一步工作进行展望.

## 2 相关工作

### 2.1 拼图问题的自动求解算法

早在 1964 年, 拼图就已经作为一种研究问题出现在学术文献中: Freeman 和 Garder<sup>[10]</sup> 提出了一种针对特定类型的无图像拼图问题 (apictorial jigsaw puzzle)、基于模式识别的自动求解方法. 后来, 拼图问题的计算复杂度得到了研究者的进一步确认: Berger<sup>[11]</sup> 证明了在一般意义上边匹配拼图问题 (edge-matching puzzle) 无法被任何自动化算法高效求解; Demaine 等<sup>[12]</sup> 证明了经典的无图像拼图、边匹配拼图, 以及组合体拼图 (polyform packing puzzle) 这 3 类问题都是 NP 完全问题.

关于拼图问题自动求解算法的研究工作大致可以分为 3 类.

第 1 类工作关注无图像拼图问题<sup>[13~16]</sup>. 其中, 无图像拼图问题通常被转化为计算几何学问题<sup>[13]</sup>. 第 2 类工作在无图像拼图问题的基础上, 为其中的每个图块附着了图像信息<sup>[17~20]</sup>. 因为图像信息的加入为拼图问题的求解提供了更多的线索, 相比较无图像拼图问题, 这类问题的困难程度有所降低. 第 3 类工作关注带图像且图块片段形状和尺寸完全相同的拼图问题<sup>[21~27]</sup> (下文简称为: 带图像拼图问题). 此类拼图问题通常被转化为计算机视觉问题<sup>[25, 26]</sup>.

大多数针对带图像拼图问题的自动求解算法通常包含两个构成成分: (1) 图块匹配度量准则, 用于对任意两个不同图块的任何一种邻接方式的匹配度进行定量评估; (2) 图块集成策略, 用于发现高质量的拼图问题候选方案. 例如, 针对图块朝向确定的带图像拼图问题的最新自动求解算法<sup>[25, 26]</sup>, 采用了由 Cho 等<sup>[21]</sup> 提出的“图块邻接边像素差值的平方和”作为图块匹配度量准则. 该算法的图块集成策略由一个遗传算法实现. 其中, 交叉操作利用了 Pomeranz 等<sup>[22]</sup> 提出的 best-buddy 策略. 又例如, 针对图块片段朝向不确定的带图像拼图问题的最新自动求解算法<sup>[27]</sup>, 采用了由 Gallagher<sup>[24]</sup> 提出的 MGC 方法作为图块匹配度量准则: 该准则利用图块边缘像素的梯度信息计算图块匹配度. 该算法的图块集成策略由一个两步算法实现: 第 1 步, 自下而上地搜索田字格图块片段组合; 第 2 步, 自上而下地对第 1 步中发现的田字格图块组合进行进一步的集成.

需要指出的是, 这些针对带图像拼图问题的自动求解算法的有效性在很大程度上依赖于以下两个假设:

- 两个图块在某种邻接方式下的匹配度正相关于这种邻接方式出现在拼图正确答案中的概率. 这一假设在一般意义上并不成立, 因为我们可以很容易地构造出一个违反这个假设的带图像拼图问题 (例如, Demaine 等<sup>[12]</sup> 的研究工作中提到的边匹配拼图问题).

- 两个图块某种邻接方式的匹配度可以基于图块的边缘像素信息计算产生. 这一假设完全不同于人类玩家的拼图求解方式: 人类玩家更依赖于两个图块上的图像内容来对图块邻接方式的匹配度进行评估. 在我们的实验中, 所有图块边缘的两圈像素被抹除为透明色, 没有任何一个人类玩家注意到这个情况, 但这种情况却使得自动化算法的求解质量产生了显著的下降.

在我们所知的范围内, 本文是第 1 个基于互联网人类群体智能进行带图像拼图问题求解的研究工作. 但是, 本文工作的目标并不是为了证明本文方法优于其他自动化的带图像拼图问题求解算法: 对于那些满足上述两个假设的拼图问题, 已有的自动求解算法已经能够进行高效高质量的求解. 如前文所述, 我们真正感兴趣是拼图问题所代表的一类复杂问题的群智化求解方法, 特别是具有复杂逻辑结构的信息制品的群智化构造方法. 我们的更远期研究目标则是希望证明: 被恰当组织的人类群体, 拥有无法轻易被机器人和算法所取代的独特能力.

## 2.2 环境激发效应与群体智能

许多研究者利用环境激发效应这一概念,解释面向人类的群体智能现象,或设计面向人类的人工群体智能系统. Parunak<sup>[28]</sup>指出,环境激发效应不是昆虫群体特有的现象:人类群体中的环境激发效应在前计算机时代以及基于计算机的社交系统中广泛存在. Parunak 进一步提出了一种针对环境激发效应的一般性架构,包含 4 个构成成分:个体状态、个体行为机理、环境状态、环境行为机理. Heylighen<sup>[29]</sup>设计了一种用于构造基于互联网的人类群体心智图 (collective mental map) 的算法. 该算法对应于 Parunak 的环境激发效应一般性架构中的环境行为机理,其输出的群体心智图则对应于该架构中的环境状态. Malone 等<sup>[30]</sup>总结了成功的互联网群体智能现象中出现的一组用于信息激发和信息片段融合的可复用模式.

我们在两个成功的人工群体智能系统中发现了本文提出的“探索-融合-反馈”回路. 一个是 UNU 系统<sup>[31,32]</sup>: 一个支持一群参与者采用协同方式求解单项选择题的在线平台. 在 UNU 系统提供的虚拟环境中,其问题求解过程可以被视为一个“探索-融合-反馈”回路: (1) 每一个参与者通过在合适的位置放置一个由其控制的磁铁来对环境中的一个铁质冰壶施加一个吸引力 (探索); (2) 所有施加在冰壶上的吸引力汇聚在一起,形成一个指向特定方向的力,并驱动冰壶在这个方向上移动 (融合); (3) 冰壶的移动行为被所有参与者实时观察到 (反馈),进而刺激每个参与者对其控制的磁铁的位置进行移动 (新一轮的探索). 另一个是 EteRNA 系统<sup>[33]</sup>: 一个支持非专业群体协同求解蛋白质结构问题的多玩家在线游戏平台. EteRNA 系统中的问题求解过程也可以被理解为一个“探索-融合-反馈”回路: (1) 玩家在各自的工作空间中设计他们的方案 (探索); (2) 玩家群体对所有个体设计的方案进行审查和投票,其中排名靠前的一组方案会通过特定的化学指标进行综合与确认 (融合); (3) 结果被在线公开发布 (反馈),进而刺激玩家进行新方案的设计 (新一轮的探索).

也有很多关于群体智能的研究工作并没有涉及环境激发效应. Levy<sup>[34]</sup>讨论了信息空间中群体智能的理想形式,并分析了信息空间中的群体智能对人类社会的影响. Woolley 等<sup>[35]</sup>通过实验展示了物理空间中面对面小规模人类群体存在群体智能现象的证据,并提出了一种量化指标对群体智能现象进行评估. Nielsen<sup>[36]</sup>认为,信息空间中群体智能的本质是一种设计导致的机缘巧合 (designed serendipity): 从一个大规模的人类群体中,更有可能发现一组人类个体,他们恰好拥有求解特定问题所需的所有信息片段. Maleszka 和 Nguyen<sup>[37]</sup>关注群体智能现象具有的“整体大于部分之和”的基本性质,提出了一种用于层次结构信息片段融合的数学模型,用于发现那些不存在于任何部分之中的新知识.

## 2.3 众包

作为一种问题求解模式<sup>[38]</sup>,众包这一概念可以从狭义和广义两个方面来理解. 在狭义上,众包指:“一个公司或组织通过公开招募的方式将一个原本由其内部雇员完成的任务外包给互联网上不可控的人类群体 (这一群体通常具有较大的规模)”;或者更为简洁,即“将任务外包给大众”<sup>[39,40]</sup>. 这是众包概念的初始含义. 在广义上,众包用于指代任何利用群体能力求解问题的方式<sup>[41]</sup>. 这种含义对众包概念的本义进行了极大的泛化,使得众包成为群体智能这一概念的同义或近义词<sup>[42]</sup>.

目前,实践中存在两种主流的众包活动. 第 1 种众包活动所关注的问题,其自身即体现为或分解为一组数量庞大的对算法而言非常困难、对人类却相对简单的子问题 (这类子问题通常被称为微任务)<sup>[43,44]</sup>. 第 2 种众包活动所关注的问题,其通常无法或很难被分解为一组相对简单的子问题. 希望求解该问题的组织在公共媒体上公开征集该问题的解决方案,并为提供最优解决方案的参与者发放具

有吸引力的奖金.

与这两种主流的众包实践活动相比, 本文方法可以被视为一种更为智能的众包, 具体表现在两个方面. 首先, 在本文方法中, 求解问题前不需将问题事先分解为一组子问题. 其次, 本文方法更多依赖于一组个体之间的自组织协同来进行问题求解, 而在当前主流的众包实践中, 几乎看不到协同现象. Kittur 等<sup>[45]</sup>认为, 协同将是未来基于群体的问题求解活动具有的一个基础性特点. 进一步而言, 如果从广义上理解众包这一概念, 则类似 UNU 和 EteRNA 这类的人工群体智能系统也可被视为一种众包系统; 在其中, 协同成为一个关键因素. 在这种意义上, 当前两种主流的众包实践活动则可被视为原始形态的人工群体智能系统.

### 3 方法

本节首先引入本文关注的一种特殊类型的拼图问题; 然后, 分别介绍基于“探索 – 融合 – 反馈”回路的群体拼图方法中探索、融合、反馈 3 个活动涉及的关键概念.

#### 3.1 拼图问题

拼图问题有很多变体. 本文关注一种特殊类型的拼图问题, 其具有如下特点:

(1) 待复原的图片 (以下简称为原图) 是一个矩形, 其内容是人类可理解的图像<sup>3)</sup>.

(2) 原图在行和列方向上被平均分割为  $M$  行和  $N$  列, 且确保分割后形成的图块是一个四边带有凹槽或突耳的正方形. 通过这种方式, 原图被分割为  $M \times N$  块大小相同且没有重叠的图块. 称该拼图问题的规模为  $M \times N$ .

(3) 所有图块的方向与其在原图中的方向一致, 且在游戏过程中不会发生变化 (即, 拼图环境不会旋转图块, 也不对玩家提供旋转操作).

(4) 玩家在拼图结束前, 无法看到原图.

给定一个规模为  $M \times N$  的拼图问题, 其正确解是由所有图块形成的一个  $M \times N$  的矩阵. 在这个矩阵中, 称相邻的两个图块 (上下相邻或左右相邻) 之间存在一个邻接关系. 可知, 在这个  $M \times N$  的矩阵中一共存在  $(2MN - M - N)$  个邻接关系: 其中, 上下邻接关系  $N(M - 1)$  个, 左右邻接关系  $M(N - 1)$  个. 因此, 求解一个规模为  $M \times N$  的拼图问题, 就等价于发现所有节点之间存在的  $(2MN - M - N)$  个邻接关系, 从而形成一个  $M \times N$  的矩阵. 称这样的一个矩阵是当前拼图问题的一个候选方案.

本文采用标签图 (labelled graph) (见图 2) 对拼图问题的候选方案进行表示. 在这种表示方式中, 节点对应于图块, 节点之间的边表示两个对应图块之间的邻接关系. 一条边上的“L-R”标签表示 L 侧的节点与 R 侧的节点在水平方向上从左到右相互邻接; 一条边上的“T-B”标签表示 T 侧的节点与 B 侧的节点在垂直方向上从上到下相互邻接. 其中, “L”, “R”, “T”, “B” 分别表示“左”、“右”、“上”、“下”的含义. 称“L”和“R”为一对相反的方向; 称“T”和“B”为一对相反的方向. 给定一个方向  $drt$ , 其反方向记为  $rev(drt)$ . 对于一条边  $(x, \text{“L-R”}, y)$ , 可以这样理解: 节点  $x$  在其中承担的角色为 L; 节点  $y$  在其中承担的角色为 R. 因此, 这种标签图本质上是一种有向图.

在不引起歧义的情况下, 下文使用拼图问题来指代上述特殊类型的拼图问题, 并使用上述基于图的术语对拼图问题及本文方法进行描述.

**定义1** (拼图问题的候选方案) 给定一个规模为  $M \times N$  的拼图问题, 其候选方案是一个图  $S$ , 满足如下条件:

3) 在实验中我们主要选取的是经典动漫影视作品相关的卡通图像.

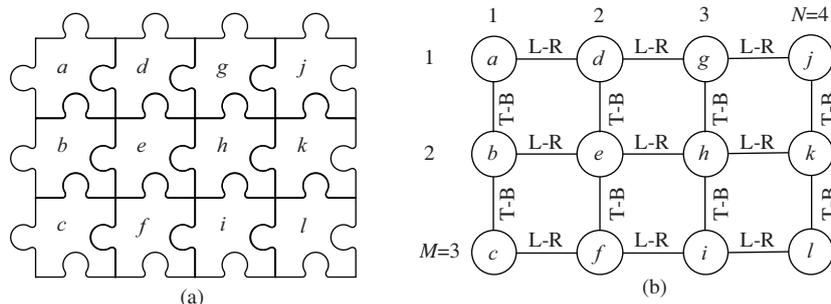


图 2 拼图问题的候选方案及其基于标签图的表示

Figure 2 A candidate solution of a PJ puzzle and its labelled graph representation. (a) A candidate-solution of a PJ puzzle of size 3×4; (b) the labelled graph representation of this candidate-solution

(1)  $|\mathcal{V}(S)| = MN$ ;

(2) 存在对  $\mathcal{V}(S)$  中所有节点的一种行列编码方式, 即:  $\mathcal{V}(S) = \{v_{ij} | 0 \leq i < M, 0 \leq j < N\}$ , 满足  $\mathcal{E}(S) = \{(v_{ij}, \text{“L-R”}, v_{i(j+1)}) | 0 \leq i < M, 0 \leq j < N-1\} \cup \{(v_{ij}, \text{“T-B”}, v_{(i+1)j}) | 0 \leq i < M-1, 0 \leq j < N\}$ .

给定一个规模为  $M \times N$  的拼图问题, 其玩家集合记为  $\mathcal{P}$ , 由所有节点构成的集合记为  $\mathcal{V}$ , 由所有候选方案构成的集合记为  $\mathcal{S}$ , 由所有候选方案中出现的边构成的集合记为  $\mathcal{E}$  (以下简称为拼图问题的边集合). 对于任何一条边  $e \in \mathcal{E}$ , 其边上的标签记为  $\text{tag}(e)$ . 对于一条标签为 “L-R” 的边  $e$ , 其中承担 “L” 角色的节点记为  $e.L$ , 承担 “R” 角色的节点记为  $e.R$ . 一条标签为 “T-B” 的边, 也用类似的方式定位其关联的两个节点. 本文正文中没有给出定义的符号, 请参见附录 A.

性质1 对于一个规模为  $M \times N$  的拼图问题: (1)  $|\mathcal{S}| = (MN)!$ ; (2)  $\forall S \in \mathcal{S} : |\mathcal{E}(S)| = 2MN - M - N$ ; (3)  $|\mathcal{E}| = 2MN(MN - 1)$ .

也即, 一个规模为  $M \times N$  的拼图问题, (1) 一共具有  $(MN)!$  个候选方案; (2) 每一个候选方案中存在  $2MN - M - N$  条边; (3) 边集合  $\mathcal{E}$  中包含  $2MN(MN - 1)$  条边. 其中, 把一个  $M \times N$  的矩阵中的所有行拼接成一行 (该行中包含  $MN$  个元素), 则可知这些元素共有  $(MN)!$  种不同的排列方式, 故性质 (1) 得证明. 性质 (2) 的证明在上文中已经给出: 即, 计算上下邻接关系和左右邻接关系数量之和. 性质 (3) 的证明类似: 在所有  $MN$  个元素, 若选出两个元素构成一条标签为 “L-R” 的边, 则共有  $MN(MN - 1)$  种不同的方式; 若选出两个元素构成一条标签为 “T-B” 的边, 则共有  $MN(MN - 1)$  种不同的方式; 则所有可能边的总数为  $2MN(MN - 1)$  条, 性质 (3) 得证.

上述 3 条性质指出了理解 “拼图问题求解” 的两种等价形式: (1) 从所有  $(MN)!$  个候选方案中找出一个正确的候选方案; (2) 从  $\mathcal{E}$  包含的  $2MN(MN - 1)$  条边中筛选出  $2MN - M - N$  条正确的边.

### 3.2 探索

在探索活动中, 每个玩家在各自的游戏页面中独立进行拼图<sup>4)</sup>. 在探索活动中, 一个玩家会执行两种类型的操作: 添加边 (即, 把不同的图块拼在一起); 删除边 (即, 把已经拼在一起的图块分开). 以标签图作为玩家拼图结果的表示方式, 则玩家在探索活动中的操作体现为: 在图的节点之间添加边, 或删除图中已经存在的边. 在探索活动的任一时刻, 对于一个玩家, 拼图环境会记录两种信息: (1) 该玩家当前时刻的拼图结果; (2) 该玩家当前时刻的拒绝边集合. 这两种信息的形式化说明分别见定义 2 和 3.

4) 不同玩家的探索活动之间并非真正独立: 在反馈活动 (见 3.4 小节) 中, 一个玩家会获得来自群体的反馈信息, 从而提高其探索活动的效率.

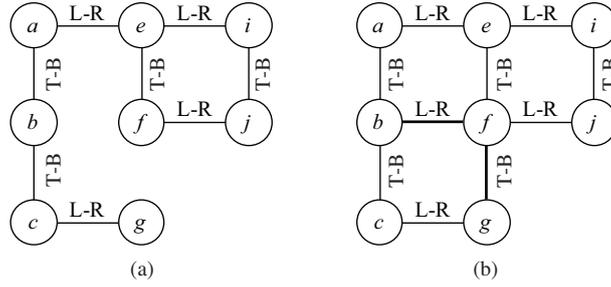


图 3 从已有边中推导出新的边

Figure 3 Deducing edges from existing edges. (a) A connected graph; (b) the connected graph adding two deduced edges

**定义 2** (一个玩家在某一时刻的拼图结果) 给定一个拼图问题及其节点集合  $\mathcal{V}$ 、边集合  $\mathcal{E}$ 、玩家集合  $\mathcal{P}$ , 一个玩家  $p \in \mathcal{P}$  在时刻  $t$  的拼图结果是一个由连通图构成的集合<sup>5)</sup>, 记为  $\mathcal{PS}(p, t)$ , 满足如下条件:

- (1)  $\forall C \in \mathcal{PS}(p, t) : \mathcal{V}(C) \subseteq \mathcal{V} \wedge \mathcal{E}(C) \subset \mathcal{E}$ ,
- (2)  $\forall v \in \mathcal{V} : \exists C \in \mathcal{PS}(p, t) : v \in \mathcal{V}(C)$ ,
- (3)  $\forall C_0, C_1 \in \mathcal{PS}(p, t) : C_0 \neq C_1 \Rightarrow \mathcal{V}(C_0) \cap \mathcal{V}(C_1) = \emptyset$ ,
- (4)  $\forall C \in \mathcal{PS}(p, t) : \forall e, f \in \mathcal{E}(C) : ((e \neq f) \wedge (\text{tag}(e) = \text{tag}(f))) \Rightarrow (((\text{tag}(e) = \text{L-R}) \wedge (e.\text{L} \neq f.\text{L}) \wedge (e.\text{R} \neq f.\text{R})) \vee ((\text{tag}(e) = \text{T-B}) \wedge (e.\text{T} \neq f.\text{T}) \wedge (e.\text{B} \neq f.\text{B})))$ ,
- (5)  $\forall C \in \mathcal{PS}(p, t) : \text{不存在由相同标签的边形成的环}$ ,
- (6)  $\forall C \in \mathcal{PS}(p, t) : \text{从 } \mathcal{E}(C) \text{ 中的边不能推导出 } \mathcal{E}(C) \text{ 中不存在的边}$ .

条件 (1) 限定, 对于  $\mathcal{PS}(p, t)$  中的任何一个连通图, 其包含的节点只能来自节点集合  $\mathcal{V}$ , 其包含的边只能来自边集合  $\mathcal{E}$ . 条件 (2) 和 (3) 限定,  $\mathcal{PS}(p, t)$  中的所有连通图对应的节点集合构成了对节点集合  $\mathcal{V}$  的一个划分. 条件 (4) 限定, 在任何时刻的拼图结果中, 任何一个图块在任何方向上, 最多只能邻接一个图块. 条件 (5) 限定, 例如, 给定 3 个节点  $i, j, k$ ,  $\mathcal{PS}(p, t)$  中的任何一个连通图, 都不能同时包含如下 3 条边:  $\langle i, \text{L-R}, j \rangle$ ,  $\langle j, \text{L-R}, k \rangle$ , 和  $\langle k, \text{L-R}, i \rangle$ . 条件 (6) 限定, 对于  $\mathcal{PS}(p, t)$  中的任何一个连通图, 从其边集中能够推导出的任何一条边一定也包含在当前边集中. 图 3 给出了边之间存在的推导关系的一个示例. 从左侧连通图的边出发, 可以推导出该连通图中缺失的两条边:  $\langle b, \text{L-R}, f \rangle$  和  $\langle f, \text{T-B}, g \rangle$ . 因此, 这两条边应该被加入到左侧连通图, 从而形成右侧的连通图.

需要指出的是, 定义 2 并不会对拼图玩家的探索活动带来任何不必要的限制: 任何一个合法的拼图结果天然会满足上述 6 种条件. 引入该定义的主要目的是确保当拼图环境对一个玩家的拼图结果进行记录或修改时, 不能违反这些条件.

为下文叙述方便, 令  $e_{\text{L}}(v, p, t)$  表示  $t$  时刻玩家  $p$  的拼图结果  $\mathcal{P}(p, t)$  中节点  $v$  在 L 方向上连接的边  $d$ . 如果这条边存在, 则可知  $e_{\text{L}}(v, p, t).\text{R} = v$ ; 如果这一条边不存在, 则  $e_{\text{L}}(v, p, t)$  返回一个 null 值. 同理, 可以定义  $e_{\text{R}}(v, p, t)$ ,  $e_{\text{T}}(v, p, t)$ ,  $e_{\text{B}}(v, p, t)$ , 分别表示节点  $v$  在 R, T, B 方向上连接的边. 令  $C(v, p, t)$  表示  $\mathcal{P}(p, t)$  中节点  $v$  所在的连通图. 令  $\mathcal{PS}$  表示由当前拼图问题所有可能的拼图结果构成的集合.

5) 在初始时刻, 一个玩家的拼图结果是一个由  $|\mathcal{V}|$  个节点和 0 条边构成的图; 可以将其视为由  $|\mathcal{V}|$  个连通图构成的集合 (其中, 每个连通图仅包含一个节点). 经过一系列的添加/删除边的操作后, 拼图结果仍然可能不是一个连通图 (即, 存在两个节点, 其之间不存在一条路径). 因此, 我们将一个玩家在任意时刻的拼图结果建模为一个由连通图构成的集合.

**定义3** (一个玩家在某一时刻的拒绝边集合) 给定一个拼图问题及其玩家集合  $\mathcal{P}$ , 一个玩家  $p \in \mathcal{P}$  在时刻  $t$  的拒绝边集合, 记为  $\mathcal{R}(p, t)$ , 定义如下:

$$\mathcal{R}(p, t) = \{e | \forall C \in \mathcal{P}(p, t) : e \notin \mathcal{E}(C), \exists u < t : \exists C \in \mathcal{P}(p, u) : e \in \mathcal{E}(C)\}.$$

也即, 一个玩家在某一时刻的拒绝边集合包含了所有那些曾经出现在该玩家拼图结果中、但没有出现在该玩家当前拼图结果中的边。

对于玩家  $p$  在时刻  $t$  的拒绝边集合  $\mathcal{R}(p, t)$  中的一条边  $e$ , 令  $R(e, t, p)$  表示  $e$  最近一次被该玩家删除时所在的连通图。

**性质2** 给定一个拼图问题及其玩家集合  $\mathcal{P}$ ,  $\forall t : \forall p \in \mathcal{P} : \forall C \in \mathcal{PS}(p, t) : \mathcal{E}(C) \cap \mathcal{R}(p, t) = \emptyset$ .

也即, 在任何时刻, 一个边不能同时包含在一个玩家的拼图结果和拒绝边集合中。

### 3.3 融合

在任何时刻, 通过融合活动, 拼图环境实时地将所有玩家的当前拼图结果和拒绝边集合融合在一起, 形成当前时刻的群体观点图。

**定义4** (某一时刻的群体观点图) 给定一个拼图问题以及玩家群体  $\mathcal{P}$ , 该玩家群体在时刻  $t$  对该拼图问题的群体观点图, 记为  $Q_t$ , 其节点集合与边集合定义如下:

- (1)  $\mathcal{V}(Q_t) = \mathcal{V}$ ,
- (2)  $\mathcal{E}(Q_t) = [\bigcup_{p \in \mathcal{P}, C \in \mathcal{PS}(p, t)} \mathcal{E}(C)] \cup [\bigcup_{p \in \mathcal{P}} \mathcal{R}(p, t)]$ .

对于时刻  $t$  的群体观点图  $Q_t$  中的每一条边  $e$ , 拼图环境会维护两个集合: 采纳这条边的玩家集合, 记为  $\mathcal{P}_{\text{sup}}(e, t)$ ; 拒绝这条边的玩家集合, 记为  $\mathcal{P}_{\text{rej}}(e, t)$ . 两者的定义如下:

- $\mathcal{P}_{\text{sup}}(e, t) = \{p | p \in \mathcal{P}, \exists C \in \mathcal{PS}(p, t) \cdot e \in \mathcal{E}(C)\}$ ,
- $\mathcal{P}_{\text{rej}}(e, t) = \{p | p \in \mathcal{P}, e \in \mathcal{R}(p, t)\}$ .

对于群体观点图  $Q_t$  中的每一条边  $e$ , 拼图环境还会维护两个权值: 正权值, 记为  $w^+(e, t)$ ; 负权值, 记为  $w^-(e, t)$ . 两者的定义如下:

- $w^+(e, t) = \sum_{p \in \mathcal{P}_{\text{sup}}(e, t), C \in \mathcal{PS}(p, t)} \mathbf{1}(e \in \mathcal{E}(C)) \cdot |\mathcal{E}(C)|$ ,
- $w^-(e, t) = \sum_{p \in \mathcal{P}_{\text{rej}}(e, t)} |\mathcal{E}(R(e, p, t))|$ .

可知, 一条边在某一个时刻的正权值, 等于当前时刻所有玩家的连通图集合中包含这条边的连通图的边数之和; 一条边在某一个时刻的负权值, 等于当前时刻所有玩家的拒绝边集合中这条边 (如果存在) 最近一次被删除时所在连通图的边数之和。

对于群体观点图  $Q_t$  中的每一条边  $e$ , 其可信系数, 记为  $\varphi(e, t)$ , 定义如下:

$$\varphi(e, t) \doteq \frac{w^+(e, t)}{w^+(e, t) + w^-(e, t)}.$$

对于群体观点图  $Q_t$  中的每一个节点  $v$ , 其在 L 方向上的边集合, 记为  $\mathcal{E}_L(v, t)$ , 定义如下:

$$\mathcal{E}_L(v, t) \doteq \{e | e \in \mathcal{E}(Q_t), e.R = v\}.$$

节点  $v$  在 L 方向上的  $\phi$ -有效边集合, 记为  $\mathcal{E}_L(v, t|\phi)$ , 定义如下:

$$\mathcal{E}_L(v, t|\phi) \doteq \{e | e \in \mathcal{E}_L(v, t), \varphi(e, t) \geq \phi\},$$

其中,  $\phi$  是位于区间  $[0, 1]$  中的一个常量值。

节点  $v$  在 L 方向上的  $(\phi, \epsilon)$  - 强有效边集合, 记为  $\mathcal{E}_L(v, t|\phi, \epsilon)$ , 定义如下:

$$\mathcal{E}_L(v, t|\phi, \epsilon) \doteq \{e|e \text{ 是 } \mathcal{E}_L(v, t|\phi) \text{ 中按正权值从大到小排序前 } k \text{ 的边, } k = \lceil [W_L(v, t|\phi)]^\epsilon \rceil\},$$

其中,  $\epsilon$  是一个接近或等于 0 的非负常量;  $W_L(v, t|\phi)$  表示由  $\mathcal{E}_L(v, t|\phi)$  中所有边的正权值降序排列形成的数列;  $\lceil (a_i)_1^k \rceil^\epsilon$  表示数列  $(a_i)_1^k$  的  $\epsilon$  - 最大差分上序列 (参见附录 A 中的定义 A1).

同理, 可定义并计算出某一时刻的群体观点图中的一个节点在 R, T, B 方向上的  $(\phi, \epsilon)$  - 强有效边集合, 分别记为:  $\mathcal{E}_R(v, t|\phi, \epsilon)$ ,  $\mathcal{E}_T(v, t|\phi, \epsilon)$ ,  $\mathcal{E}_B(v, t|\phi, \epsilon)$ .

$(\phi, \epsilon)$  - 强有效边集合包含从当前时刻的群体观点图包含的边中过滤出那些最有可能出现在正确答案中的边. 这些边将会通过反馈活动在恰当的时机反馈给群体中的特定玩家.

### 3.4 反馈

在反馈活动中, 对每一个玩家, 拼图环境会根据当前时刻的群体观点图和玩家的当前拼图结果, 向其展示特定的反馈信息, 以加快其拼图进程. 反馈活动涉及 3 个基本方面: 反馈的时机 (when, 什么时候反馈)、反馈的内容 (what, 反馈什么信息)、反馈的形式 (how, 以什么样的形式表现出反馈信息).

#### 3.4.1 反馈的时机、内容及形式

我们区分了两种类型的反馈时机: 操作后、停滞时. 所谓操作后, 就是玩家完成一次操作后的那个时刻点. 所谓停滞时, 就是玩家在上一操作完成后经过了一定的时间间隔仍然没有实施下一次操作的那个时刻点 (此时, 认为玩家的拼图活动进入了停滞状态).

从构成成分上看, 反馈的内容就是一条或多条不存在于当前玩家当前拼图结果中的边; 这些被反馈的边出自某个节点在某个方向的  $(\phi, \epsilon)$  - 强有效边集合.

反馈内容中的每一条边, 存在两种可能的反馈形式: 拼接形式、提示形式.

(1) 拼接形式: 在玩家的拼图页面上按照待反馈边的语义将对应的两个图块拼接在一起. 在拼图页面上进行这种拼接操作前, 拼图环境要确保在玩家的当前拼图结果中添加待反馈的边后 (如有必要, 还需要添加通过推导产生的边, 合并相关的连通图), 玩家的当前拼图结果能够满足定义 2 对拼图结果的要求; 否则, 说明针对当前待反馈边的拼接反馈会导致冲突.

(2) 提示形式: 在玩家的拼图页面上按照待反馈边的语义在对应的两个图块上显示特定的提示信息, 提示玩家这是一条可能正确的边. 在我们已经实现的拼图环境中, 这种提示信息表现为, 将涉及的两个图块的两条相邻边线用同一种颜色高亮显示出来.

对于玩家而言, 如果一条被反馈的边是正确的, 且这条边是通过拼接形式进行反馈的, 则玩家不需要进行任何操作; 如果这条边是通过提示形式反馈的, 则玩家需要手工将两个相关的图块拼接在一起. 如果一条被反馈的边是错误的, 且这条边是通过提示形式反馈的, 则玩家只需要忽略提示信息即可; 如果这条边是通过拼接形式进行反馈的, 则玩家需要手工拆开相关的图块. 可以看到, 对于正确的反馈, 不会对玩家造成负面影响; 对于提示形式的错误反馈, 会对玩家造成轻微的负面影响; 对于拼接形式的错误反馈, 会对玩家的探索活动带来干扰. 为了避免这种干扰现象的频繁发生, 需要尽可能提高反馈信息的准确率.

基于对反馈时机、反馈内容和反馈形式的不同组合, 我们设计了两种互补的反馈策略: 反应式反馈、刺激式反馈.

### 3.4.2 反应式反馈

在反应式 (responsive) 反馈策略中, 当玩家完成一个操作后, 拼图环境会针对该操作的焦点连通图, 根据当前时刻的群体观点图和玩家的当前拼图结果, 为玩家提供即时、局部性的反馈信息. 其反馈内容和反馈形式通过下述流程确定. 第 1 步, 获得玩家当前操作的焦点连通图. 第 2 步, 对于焦点连通图中每一个度小于 4 的节点, 对于该节点的每一个没有边连接的方向, 检查节点在该方向的  $(\phi, \epsilon)$ -强有效边集合: 如果集合中只存在一条边, 且将这条边加入玩家的当前拼图结果并添加必要的推导边以及合并连通图后, 玩家的拼图结果符合定义 2 对拼图结果的要求, 则将这条边通过拼接形式反馈给玩家. 第 3 步, 如果经过第 2 步后, 没有发生任何反馈操作, 则随机选择一个遍历过的  $(\phi, \epsilon)$ -强有效边集合, 将该集合中的所有边通过提示形式反馈给玩家.

### 3.4.3 刺激式反馈

在刺激式 (stimulative) 反馈中, 当玩家进入停滞状态后, 拼图环境会根据当前时刻的群体观点图和玩家的当前拼图结果, 主动为玩家提供相关信息的反馈, 以刺激玩家采取进一步的操作. 其反馈内容和反馈形式通过下述流程确定. 第 1 步, 从当前时刻的群体观点图中定位到所有同时满足如下 2 个条件的边: (1) 这条边没有出现在玩家当前时刻的拼图结果中; (2) 这条边连接的两个节点在对应方向上的  $(\phi, \epsilon)$ -强有效边集合只包含这条边 (例如, 对于边  $\langle u, L-R, v \rangle$ , 这表明  $\mathcal{E}_R(u, t | \phi, \epsilon) = \mathcal{E}_L(v, t | \phi, \epsilon) = \{\langle u, L-R, v \rangle\}$ ). 第 2 步, 以随机方式遍历第 1 步定位到的每一条边: 对于这样一条边, 如果将其加入玩家当前时刻的拼图结果并添加必要的推导边以及合并连通图后, 玩家的拼图结果符合定义 2 对拼图结果的要求, 则将这条边通过拼接形式反馈给玩家. 第 3 步, 如果经过第 2 步后, 没有发生任何反馈操作, 则随机选择第 1 步定位到的一条边, 将这条边通过提示形式反馈给玩家.

## 4 实验与评估

我们通过一组受控实验对本文方法的效果进行了初步的评估. 本节首先介绍实验关注的 3 个研究问题; 然后, 分别对实验平台、实验方法、实验任务与设计进行说明; 最后, 结合 3 个研究问题对实验结果进行分析.

### 4.1 研究问题

实验主要关注如下 3 个问题.

- **RQ1: 群体性能.**

以拼图问题规模和玩家群体规模为两个独立的自变量、以群体性能为因变量, 建立自变量与因变量之间的经验型因果关系.

- **RQ2: 反馈准确率与反馈比例.**

观察在不同的拼图问题规模和玩家群体规模的组合下, 群体中 fastest 完成拼图的玩家接受到的反馈信息的准确率、以及反馈信息在此类玩家拼图结果中所占的比例, 验证本文提出的融合与反馈算法的有效性.

- **RQ3: 基于环境激发效应的协同与面对面协同以及拼图问题自动求解算法的对比.**

将本文基于环境激发效应的群体拼图方法与基于面对面协同的群体拼图方法、以及拼图问题自动求解算法进行对比, 考察本文方法的优缺点.

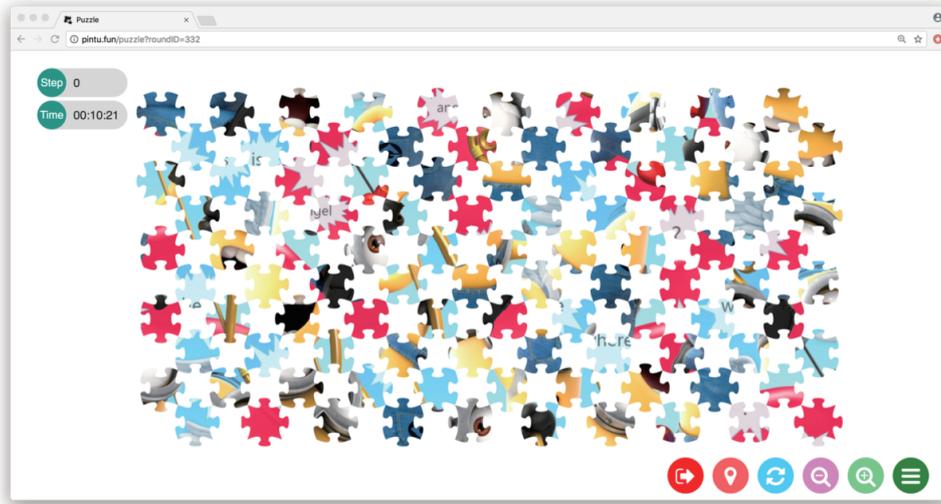


图 4 一个玩家求解拼图问题的工作空间的屏幕截图  
Figure 4 Screenshot of a player's puzzle-solving workspace

## 4.2 实验平台

我们开发了一个支持本文方法的多玩家在线拼图环境<sup>6)</sup>。实验所在 web 服务器的基本配置为: CentOS 7 操作系统、4 核 CPU、8 GB 内存。实验内容由一组拼图游戏构成: 每一次拼图游戏涉及一个拼图问题和一个玩家群体; 每一个玩家在自己的工作空间中进行拼图问题的求解。图 4 展示了一个玩家在一次拼图游戏开始时在自己的浏览器页面中看到的工作空间<sup>7)</sup>。

## 4.3 实验方法

### 4.3.1 参与者

我们通过大学校园 BBS 招募到了 52 名参与者: 年龄范围 19~45 岁, 平均年龄 34.23 岁, 年龄标准偏差 4.68 岁, 男性 30 名。其中, 32 人为博士或硕士研究生, 占比 61.5%; 15 人为本科生, 占比 38.9%; 5 人为大学教工, 占比 9.6%。我们使用酬金作为激励因素。在每一次拼图游戏中, 支付给参与者的酬金分为两个部分: 基础酬金、奖励酬金。具体而言, 在一次有  $n$  个人参与的拼图游戏中, 基础酬金和奖励酬金的总和是  $5n$  元人民币。其中,  $2.5n$  元作为基础酬金平均发放给每个参与者。剩下的  $2.5n$  元则作为奖励酬金发放给完成度排名前 50% 的参与者; 每一个参与者获得的奖励酬金正比于其拼图完成度; 一个参与者的拼图完成度的计算方式为游戏结束时 (任何一个参与者完成拼图, 则游戏结束) 该参与者的拼图结果中的正确边的数量除以  $2MN - M - N$ ; 如前文所述,  $M \times N$  为当前拼图问题的规模。

### 4.3.2 实验过程

在参与实验前, 参与者在拼图环境中注册一个账号, 学习拼图环境的基本使用方式, 并至少完成一次单人拼图游戏。在实验过程中, 每一个参与者按顺序进行如下 3 个活动: (1) 登录系统, 加入一个拼图游戏。(2) 对于一个拼图游戏, 当加入的玩家达到预定的人数后, 拼图游戏正式开始。(3) 当任何一

6) <http://www.pintu.fun>.

7) <http://www.pintu.fun/demos> 提供了两个群体拼图的视频。一个视频展示了一个问题规模为  $10 \times 10$ 、群体规模为 10 的拼图游戏中 4 个玩家的并行拼图过程。另一个视频对群体拼图游戏中一个玩家的拼图活动进行了注解说明。

表 1 350 次拼图游戏的批次编号  
Table 1 The 7 batches of 350 game rounds

ps	gs									
	1	2	3	4	5	6	7	8	9	10
4×4	7	7	7	7	6	5	4	3	2	1
5×5	6	6	6	6	6	5	4	3	2	1
6×6	5	5	5	5	5	5	4	3	2	1
7×7	4	4	4	4	4	4	4	3	2	1
8×8	3	3	3	3	3	3	3	3	2	1
9×9	2	2	2	2	2	2	2	2	2	1
10×10	1	1	1	1	1	1	1	1	1	1

个玩家完成拼图后, 所有玩家停止拼图, 填写关于这次拼图游戏的问卷, 然后退出当前游戏. 其中, 问卷的核心问题是一个关于拼图过程中接收到的反馈信息的有效性的打分评估. 该问题的可选打分包括 0.5, 1, 1.5, 2, 2.5, 3, 3.5, 4, 4.5, 5 等 10 个分数. 其中, 前 2 个分数对应的说明为 “No help at all”; 后面 8 个分数对应的说明分别为 “Better than nothing”, “Little help”, “Just so so”, “Need improvement”, “Useful help”, “Reasonable help”, “Great help”, “Significant help”.

#### 4.3.3 参数设置

实验包含的每一个拼图问题, 其所有图块 (四边带有凹槽或凸耳的正方形) 都具有相同的尺寸; 每一个图块最外围的两行/列像素的颜色被抹除 (即, 被设置为透明色). 拼图问题的规模 (ps) 变化范围从  $4 \times 4$  到  $10 \times 10$ . 玩家群体的规模 (gs) 变化范围从 1 到 10. 因此, 一共具有 70 种不同的 ps 与 gs 的组合方式. 我们预先选择了 80 多张不同的图片作为拼图问题的原图. 参数  $\phi$  和  $\epsilon$  (见第 3.3 小节) 的值分别被设为 0.618 和 0.02.

### 4.4 实验任务与设计

#### 4.4.1 RQ1: 群体性能

问题规模 ps 和群体规模 gs 被作为两个自变量, 群体性能 cp 被作为因变量. 其中, 群体性能 cp 定义为玩家群体中第一个完成拼图的玩家所使用的时间 (以秒为单位). 对每一种 ps 与 gs 的组合, 随机选择相应数量的 5 个玩家群体, 进行 5 次不同的拼图游戏. 总计进行 350 次拼图游戏.

由于实验人数以及实验成本的限制, 我们没有严格检测并禁止 “同一个玩家参与的两次拼图游戏具有相同的原图” 这种现象, 但对这 350 次拼图游戏的实验顺序进行了约束. 具体实验设计如下. 首先, 这 350 次拼图游戏被分为 7 批 (见表 1), 编号分别为  $1, 2, \dots, 7$ . 然后, 按照编号从小到大的顺序, 依次进行这 7 批游戏. 其中, 对于第  $i \in [1, 7]$  批游戏, 可知其包含  $5(18 - 2i)$  次拼图游戏, 因此, 我们随机从原图库中选择  $5(18 - 2i)$  张不同的图片, 分别作为每次拼图游戏的原图. 对于具有相同 ps 与 gs 组合的 5 次拼图游戏, 采用无放回抽样的方式随机选择 5 个玩家群体, 分别作为每次拼图游戏的玩家群体.

这种实验设计是为了尽可能消除玩家对图片的无意识记忆对群体性能 cp 造成的夸大. 如果一个玩家在一次单人拼图游戏中成功还原了一幅原图, 然后该玩家又参与到一个具有相同原图的 10 人拼图游戏中, 这时, 该玩家对原图的无意识记忆可能会加速其拼图过程, 从而使得这次 10 人拼图游戏的

群体性能比无记忆情况下的群体性能要好, 进而夸大了群体规模对群体性能造成的影响. 上述实验设计确保了如果一个图片出现在两次不同的拼图游戏中, 那么, 前一次拼图游戏的群体规模一定大于后一次拼图游戏的群体规模. 因此, 可以在很大概率上保证, 通过实验观察到的群体规模对群体性能的影响是对真实影响的一个保守观察.

#### 4.4.2 RQ2: 反馈准确率与反馈比例

我们设计了两种量化指标对反馈准确率与反馈比例进行度量.

- 反馈准确率  $f_p$ . 给定一次拼图游戏, 将最快完成拼图的玩家接收到的所有反馈信息中正确信息的比例, 作为这次游戏中反馈准确率的量化度量指标.

- 反馈比例  $f_r$ . 给定一次拼图游戏, 将最快完成拼图的玩家的拼图结果中由反馈提供的信息所占的比例, 作为这次游戏中反馈比例的量化度量指标.

我们认为, 对于最快完成拼图的玩家而言, 反馈准确率和反馈比例越高, 则表明群体的协同效果越好. 反馈准确率越高, 表明通过综合每一个玩家的当前拼图结果形成的强有效边集合的准确率越高. 反馈比例越高, 表明最快完成拼图的玩家从群体中其他玩家获得的帮助越多. 如果考察其他玩家 (例如, 最慢完成拼图的玩家) 的这两个指标, 我们认为, 在统计意义上, 其值应该比最快完成拼图玩家的这两个指标值要高, 因为反馈信息的来源是强有效边集合, 且随着时间的推移, 强有效边集合中的元素越多、元素的正确率也越高.

除了上述定量度量方式外, 我们还通过问卷调查的方式让玩家在退出拼图游戏前对其在游戏过程中感受到的反馈信息的质量进行定性评估.

#### 4.4.3 RQ3: 基于环境激发效应的协同与面对面协同以及拼图问题自动求解算法的对比

为了将本文方法与传统面对面协同拼图方式的效果进行对比, 我们进行了另外一组多人面对面的拼图游戏. 我们通过无放回抽样的方式, 确定了 5 个没有交集的玩家群体, 每个玩家群体中有 10 名参与者. 每一个玩家群体进行 7 次拼图游戏; 其中, 每一次拼图游戏的原图都不相同,  $p_s$  与  $g_s$  的组合也不相同. 在每一次拼图游戏中, 10 名玩家围坐在一个投影屏幕前, 屏幕上显示的是拼图环境中的单人工作空间; 其中, 一个玩家负责操作拼图, 其他玩家自由发表自己的拼图决策.

为了将本文方法与拼图问题自动求解算法进行对比, 我们选择针对图块片段朝向确定的带图像拼图问题的最新自动求解算法<sup>[25,26]</sup>, 让其求解 4.4.1 小节中 10 人拼图游戏中的拼图问题. 该求解器运行在一个台式计算机上, 基本配置为: CPU 3.6 GHz Intel Core i7、内存 16 GB、操作系统 Ubuntu 18.04.

### 4.5 实验结果与分析

#### 4.5.1 RQ1: 群体性能

图 5 展示了不同规模的玩家群体在求解不同规模的拼图问题时所使用的时间. 其中, 玩家群体规模  $g_s$  的变化区间为  $[1, 10]$ ; 拼图问题规模  $p_s$  的变化区间为  $[4 \times 4, 10 \times 10]$ ; 每一种  $g_s$  与  $p_s$  组合对应的时间是 5 次拼图游戏时间的平均值. 从实验结果可以观察到如下两点一般性结论: (1) 在拼图问题规模确定的情况下, 拼图问题求解时间随着玩家群体规模的增加而逐渐减少; (2) 在玩家群体规模确定的情况下, 拼图问题求解时间随着拼图问题规模的增加而逐渐增多.

表 2 展示了与单玩家相比, 玩家以群体方式进行拼图时, 拼图时间的降低比例. 如 4.4.1 小节所述, 对每一种  $p_s$  与  $g_s$  的组合, 我们随机选择了 5 个玩家群体, 进行 5 次不同的拼图游戏. 其中, 第 1 列展示了 5 个不同的玩家独立求解一个特定规模的拼图问题时, 最快完成玩家的游戏时间 (以下简称为:

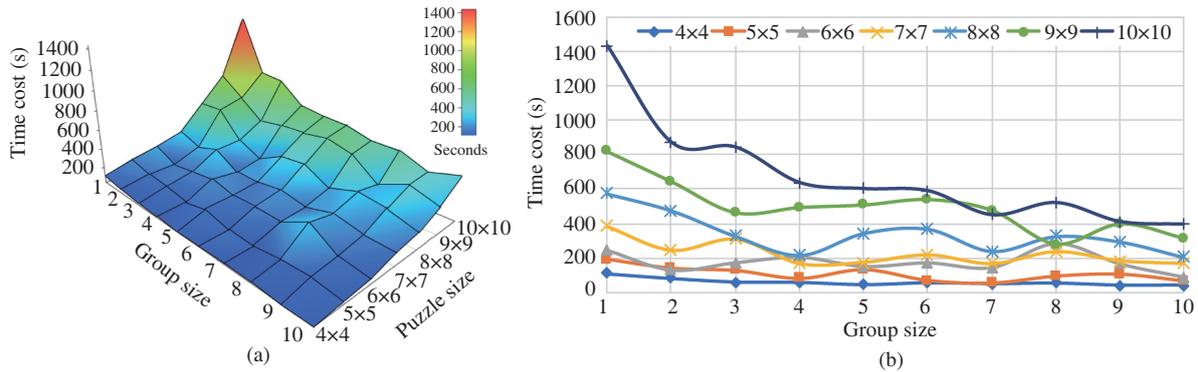


图 5 不同规模的群体  $gs \in [1, 10]$  求解不同规模的拼图问题  $ps \in [4 \times 4, 10 \times 10]$  使用的平均时间

Figure 5 The average time for player groups with  $gs \in [1, 10]$  to solve PJ puzzle with  $ps \in [4 \times 4, 10 \times 10]$ . (a) A 3D view; (b) a 2D view

表 2 群体拼图相比较单玩家拼图在求解时间上的提升

Table 2 Time improvement brought by playing as a group, comparing with the best single players

ps	gs									
	1 (s)	2 (%)	3 (%)	4 (%)	5 (%)	6 (%)	7 (%)	8 (%)	9 (%)	10 (%)
4×4	108.12	26.93	47.51	50.05	63.01	51.91	56.99	53.76	65.78	64.86
5×5	191.50	26.89	34.20	58.23	31.59	64.49	72.32	51.18	45.43	65.83
6×6	244.67	47.28	29.56	16.83	38.49	30.11	41.14	16.49	32.56	64.50
7×7	385.38	36.69	19.50	57.44	55.37	43.95	57.44	38.76	53.03	57.23
8×8	575.18	20.92	43.32	62.62	40.31	36.02	58.62	43.50	49.06	64.88
9×9	821.17	21.58	43.80	40.07	38.25	34.35	42.14	66.26	51.89	62.16
10×10	1432.12	39.20	41.17	55.56	58.03	58.90	68.68	63.72	71.44	72.51
Average	536.88	31.36	37.01	48.69	46.44	45.68	56.76	47.67	52.74	64.57

单玩家最快游戏时间). 例如, 第 1 列第 4 行中的数字 385.38 表示了 在  $ps = 7 \times 7, gs = 1$  的 5 次拼图游戏中, 单玩家最快游戏时间是 385.38 s. 第  $i \in [2, 10]$  列展示了在  $gs = i$  的 5 次拼图游戏中, 平均游戏时间相对于单玩家最快游戏时间 (即, 所在行第 1 列中的时间值) 减少的百分比. 每一列的最后一行表示了所在列中相关数据的平均值. 从这些数据中, 可以观察到, 在群体规模变化区间为  $[2, 10]$  的情况下, 在平均意义上, 相对于单玩家拼图, 群体拼图带来的拼图性能的提升范围在 31.36% 和 64.57% 之间. 基于前文所述的实验设计可知, 这一通过实验得到的性能提升范围是对真实性能提升范围的一个保守估计.

图 6 展示了在求解  $10 \times 10$  的拼图问题时, 对于不同规模的群体, 其拼图进展程度随时间的变化规律. 在这个图中, 10 人群体拼图时间的平均值被归一化为 100. 群体在某个时刻上的拼图进展程度被定义为: 当前时刻群体观点图中所有节点在所有方向上的  $(\phi, \epsilon)$ -强有效边集合中包含的正确边的数量占正确答案中边数量的百分比. 从这 10 条进展程度曲线中, 可以观察到如下 2 个现象: (1) 在进展程度区间  $[0, 0.85]$  上, 进展程度随时间变化的规律基本上是一个线性函数, 且随着群体规模的增加, 线性函数的斜率逐渐增加 (也即, 问题求解的效率逐渐提高); (2) 当群体规模从 1 增加到 2 或 3 时, 线性函数斜率的增幅最大, 这表明即使对于只有 2 或 3 个玩家的小规模群体, 当他/她们协同进行拼图

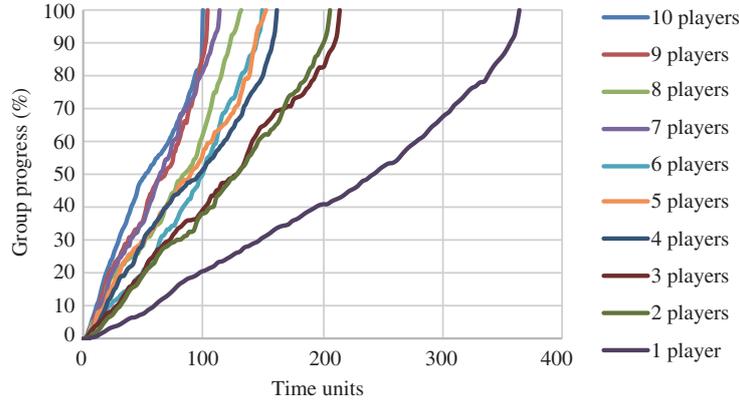


图 6 不同规模的群体  $gs \in [1, 10]$  求解  $10 \times 10$  拼图问题的时间 - 进度曲线

Figure 6 The puzzle-solving progress of a player group with  $gs \in [1, 10]$  to solve a  $10 \times 10$  PJ puzzle

时, 拼图效率就会带来显著的提升 (在我们的实验中, 拼图效率几乎提高了一倍).

为了更为准确地刻画群体性能  $cp$  与群体规模  $gs$ 、问题规模  $ps$  这两个因素之间的关系, 基于对着这 3 个因素的理解, 我们设计了如下 3 种形式的函数来刻画这 3 个因素之间的关系.

$$(1) \quad cp = f(ps, gs \mid \mu, \nu) = \mu \cdot e^{\nu \cdot ps} \cdot gs^{-1},$$

$$(2) \quad cp = f(ps, gs \mid \mu, \nu, \omega) = \mu \cdot e^{\nu \cdot ps} \cdot (gs + \omega)^{-1},$$

$$(3) \quad cp = f(ps, gs \mid \mu, \nu, \omega) = \mu \cdot e^{\nu \cdot ps} \cdot e^{-\omega \cdot gs}.$$

其中, 公式项  $\mu \cdot e^{\nu \cdot ps}$  来源于“拼图问题在一般意义上是一个 NP 完全问题”这一事实; 因此, 在群体规模确定的情况下, 拼图时间将随着问题规模的增大而指数增长. 公式项  $gs^{-1}$  反应了在线性协同 ( $1+1=2$ ) 的情况下, 拼图时间随群体规模变化的规律: 如果一个玩家独立拼图所需的时间为  $t$ , 那么, 在线性协同的情况下,  $n$  个玩家拼图所需的时间应该为  $t/n$ . 公式项  $(gs + \omega)^{-1}$  在  $gs^{-1}$  中底数  $gs$  上增加了一个常数项, 可以用于表示在亚线性协同 ( $1+1 < 2$ , 此时应有  $\omega > 0$ ) 或超线性协同 ( $1+1 > 2$ , 此时应有  $\omega < 0$ ) 情况下, 拼图时间随群体规模变化的规律. 公式项  $e^{-\omega \cdot gs}$  建模了一种超线性协同的情况: 随着群体规模的增加, 拼图时间呈指数下降趋势.

基于这 3 种函数, 我们使用线性回归技术对实验数据进行了拟合, 得到了如下 3 个具体函数, 以及对应的相关系数  $r^2$ .

$$(1) \quad 39.661 \cdot e^{0.381 \cdot ps} \cdot gs^{-1}, \quad r^2 = 0.6417,$$

$$(2) \quad 149.50 \cdot e^{0.362 \cdot ps} \cdot (gs + 3.391)^{-1}, \quad r^2 = 0.8982,$$

$$(3) \quad 36.307 \cdot e^{0.361 \cdot ps} \cdot e^{-0.130 \cdot gs}, \quad r^2 = 0.8893.$$

基于相关系数, 我们认为, 后两个函数都对实验数据进行了很好的拟合. 但是, 这两个函数的物理意义却发生了冲突. 基于第 2 个函数, 可以得出结论: 在实验中, 2~10 人的群体中发生了亚线性协同现象. 基于第 3 个函数, 则可以得出结论: 在实验中, 2~10 人的群体中发生了超线性协同现象. 我们认为, 这种冲突产生的主要原因在于实验中群体规模的变化范围还不够大. 后续研究会探索更大规模群体在求解拼图问题上的性能表现.

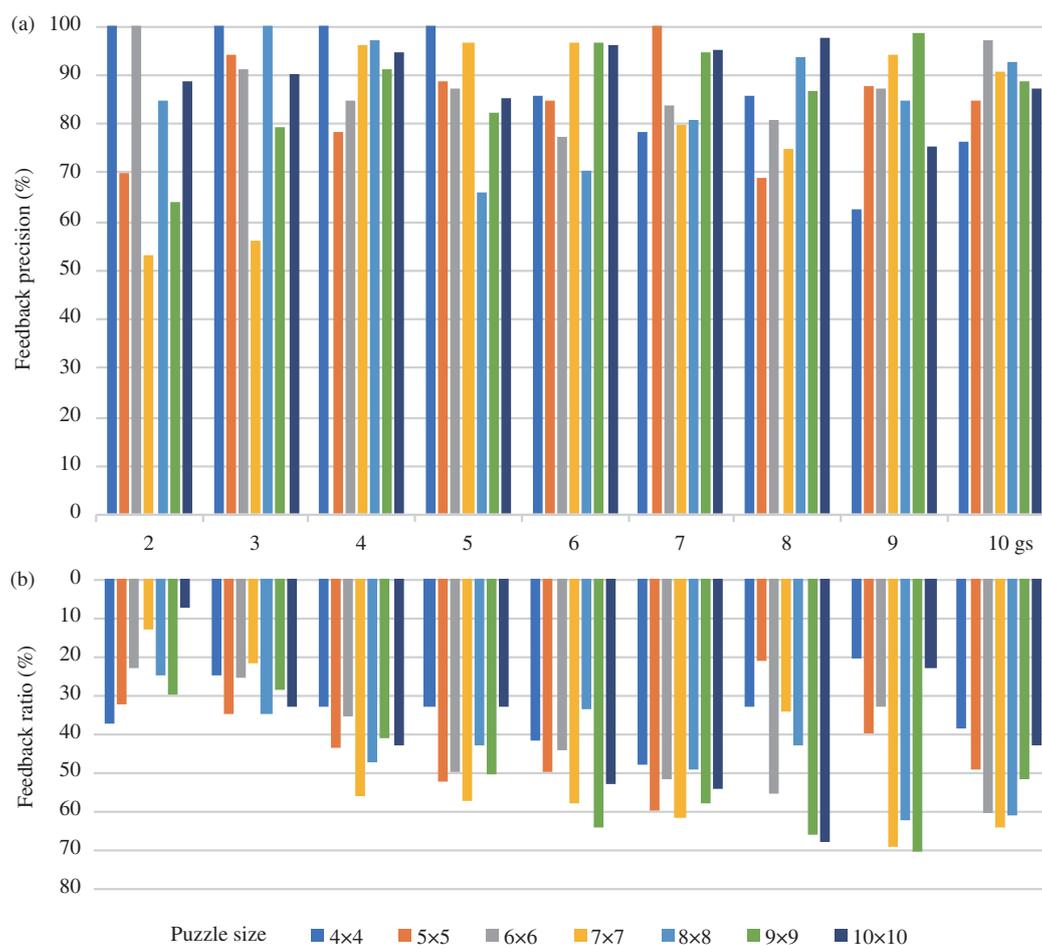


图 7 不同问题规模与群体规模组合对应的平均反馈准确率和反馈比例

Figure 7 Average feedback precision (a) and feedback ratio (b) for different combinations of puzzle size and group size

#### 4.5.2 RQ2: 反馈准确率与反馈比例

图 7 展示了在不同的问题规模  $ps$  和群体规模  $gs$  的组合下, 5 次拼图游戏的平均反馈准确率  $fp$  和反馈比例  $fr$ . 可以看到, 在实验中, 平均反馈准确率  $fp$  的均值为 86.34%, 标准偏差为 11.22%. 这表明本文方法中的融合与反馈机制具有很好的准确率和稳定性. 另外, 需要指出的是, 因为反馈的信息全部来源于在融合活动中计算产生的  $(\phi, \epsilon)$ -强有效边集合, 这基本上也就表明强有效边集合中元素的准确率也在 86.34% 左右.

同时, 可以观察到, 随着群体规模  $gs$  的增加, 平均反馈比例  $fr$  呈现出逐渐增加的趋势. 在群体规模区间  $[8, 10]$  上, 平均反馈比例的取值在 45% 左右; 这意味着, 在最快完成拼图的玩家给出的答案中, 有 45% 左右的信息是通过群体反馈获得的. 不过, 相比较群体规模区间  $[4, 7]$ , 在区间  $[8, 10]$  上, 不同问题规模下的平均反馈比例表现出较大的标准偏差. 目前我们还无法对这一现象给出合理的解释.

图 8 展示了玩家对其拼图过程中接收到的反馈信息的定性评估结果: 其中, 64% 的玩家给出的分数在 3.5 分以上 (满分为 5 分), 与图 7 中的定量评估结果基本吻合. 这也从一个侧面反映出如下信息: 在本文方法能够达到的平均反馈准确率下, 64% 的玩家能够容忍错误的反馈信息对其造成的干扰.

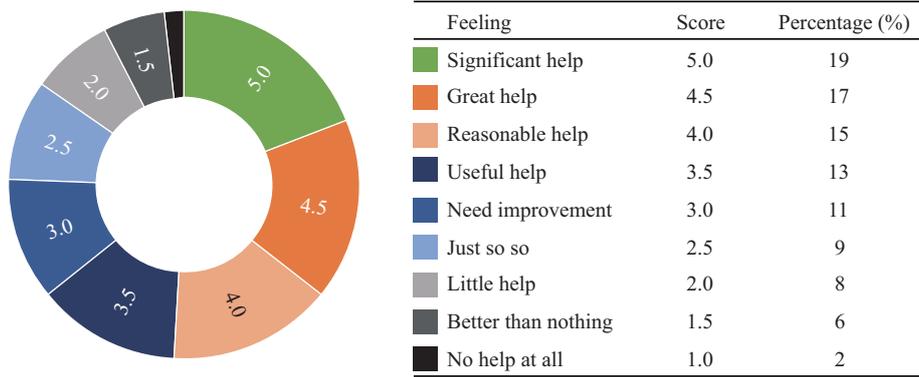


图 8 玩家对其在拼图过程中接收到的反馈信息的定性评估

Figure 8 Qualitative evaluations from players about the feedback information received in PJ puzzle solving

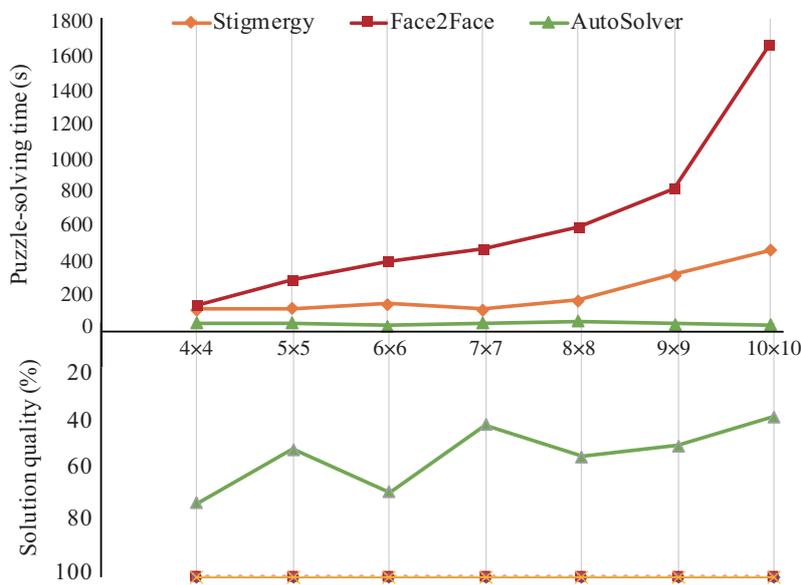


图 9 基于环境激发效应的协同、面对面协同、自动求解算法三者 in 拼图求解时间和求解质量上的对比

Figure 9 Puzzle-solving time and quality of stigmergy-based collaboration, face-to-face collaboration, and auto-solver

#### 4.5.3 RQ3: 基于环境激发效应的协同与面对面协同以及拼图问题自动求解算法的对比

图 9 展示了本文方法与面对面协同拼图方式以及拼图问题自动求解算法<sup>[25,26]</sup>在问题求解时间和求解质量两个指标上的对比. 这里, 求解质量定义为: 在最终输出的拼图问题候选方案中, 正确边的数量占正确边总数的百分比. 在这 3 种不同的方式中, 拼图问题自动求解算法具有最少的问题求解时间: 对于  $10 \times 10$  的拼图问题, 只需要 11 s 就能确定一个候选方案. 但是, 拼图问题自动求解算法却表现出较差的问题求解质量: 其求解质量的均值为 0.52, 而基于环境激发效应的协同与面对面协同的求解质量均为 1. 在两种协同式拼图方式中, 拼图问题求解时间都随着问题规模的增加而增长. 但是, 面对面协同的拼图, 在求解时间的增长速度显著大于基于环境激发效应的拼图. 也就是说, 对于 10 人群体, 相比较面对面协同的拼图, 本文方法展现出更好的群体规模可扩展性.

## 5 讨论

本节简要讨论本文研究工作涉及的几个基本问题。

影响实验结论有效性的风险因素. 作为一项初步研究, 在本文工作中, 我们仅招募了 52 名志愿者参与到实验中. 参与者大多数为作者所在高校的学生, 缺乏足够的多样性. 本文采取了一些策略来尽量减少实验结果中可能存在的偏差. 例如, 对每一种问题规模与群体规模的组合, 进行 5 次不同的拼图游戏, 取其平均值作为最终实验结果数据; 设计了一种分批执行的实验方案, 避免玩家对图片的无意识记忆对群体性能的激进影响. 即便如此, 我们也无法确保实验结果中不存在任何的统计性偏差. 提高实验结果准确性的更可靠手段是在更大规模的问题和玩家群体上进行更加全面的实验; 这将是后续的一项重要研究工作.

玩家观点存在冲突对群体拼图性能的影响. 虽然本文实验中涉及的拼图问题只具有唯一的正确答案, 但在拼图的过程中, 我们确实观察到不同的玩家在某些块之间的邻接关系上存在暂时性的冲突. 基于本文方法的特点, 我们认为, 对于存在冲突的不同观点, 只要持正确观点玩家的数量占有所有持有观点的玩家总数的比例大于 0.618 (即, 本文实验中  $\phi$  的取值), 就不会对强有效边集合的计算结果产生显著的影响, 进而不会对反馈准确率造成显著影响. 同时, 由于只存在唯一的正确答案, 即便对于那些持有错误观点的玩家, 随着游戏的持续进行, 这些玩家会逐步发现并改正其持有的错误观点, 从而回到正确的轨道上来. 因此, 在很大意义上, 本文的实验结果已经体现了玩家观点存在暂时性冲突对群体拼图性能的影响. 另一方面, 对于那些不存在唯一正确答案的复杂问题, 在存在观点冲突的情况下, 一种合理的应对方案是: 把这些存在冲突的观点反馈给群体, 让群体中的每一个体独立判断其所赞同的观点, 并在此基础上进行各自的探索活动. 对不存在唯一正确答案的复杂问题的群智化求解, 也将是我们后续的一项重要研究工作.

群体规模与问题规模的可扩展性. 基于环境激发效应的群体智能具有良好的群体规模可扩展性. 但在技术层次上, 本文方法的群体规模可扩展性依赖于一个重要因素: 群体拼图环境所在的服务器需要有足够的计算能力, 能够低延时地响应用户的操作 (对其进行反馈). 在自然界基于环境激发效应的群体智能现象中, 问题规模的可扩展性并不是一个重要问题, 因为社会性昆虫群体面临的问题通常具有相对固定的问题规模. 本文方法的问题规模可扩展性依赖于显示拼图空间的显示器的大小和分辨率. 同时, 在实验中, 我们也观察到, 如果拼图问题的规模太大, 会导致玩家在显示器上看到过多的图块, 使得玩家产生消极情绪. 这也是影响本文方法问题规模可扩展性的一个重要因素. 后续工作将寻求相应的可能解决方案 (例如, 仅为一个玩家随机展示一定比例的图块).

群体智能的定量评估. 本文提出了一种基于群体智能的拼图问题求解方法. 但对方法的评估目前还停留在定性层次上. 我们认为, 对人类群体智能系统的定量评估可以从该系统的群体规模可扩展性和问题规模可扩展性两个方面进行. 给定两个求解同类问题的群体智能系统, 可以根据“问题规模确定情况下, 两者在问题求解时间随群体规模增加的降低速度”, 对两者在群体规模可扩展性方面的差异进行定量评估; 可以根据“群体规模确定情况下, 两者在问题求解时间随问题规模增加的增长速度”, 对两者在问题规模可扩展性方面的差异进行定量评估. 在此基础上, 可以对这两个系统的群体智能程度的相对高低进行定量评估.

“探索 - 融合 - 反馈”回路对其他问题的适用性. “探索 - 融合 - 反馈”回路明确了环境激发效应中环境承担的两种职责 (即, 融合与反馈), 进而指出了一种在信息空间中构造人工群体智能系统的一般性技术框架. 我们认为, “探索 - 融合 - 反馈”回路能否适用于特定复杂问题的求解, 主要依赖于能否有效地对群体中不同个体拥有的问题相关信息进行有效地激发、表示、融合与反馈. 一个重要的研

究问题是, 能否找到一般性的独立于问题的信息表示、融合与反馈机制. 进一步而言, 本文的研究工作是本文作者之一在 2017 年发表的一篇关于“基于群体智能的软件开发”的观点性论文<sup>[5]</sup>的探索性尝试. 这篇观点性论文在对“软件开发没有银弹”这一主观判断进行分析的基础上, 以自然界和互联网中的群体智能现象为参照物, 从“软件开发的本质性困难”、“软件开发组织方式的演化趋势”、“软件制品的结构特点”3 个方面对基于群体智能的软件的可行性进行了分析, 总结了 3 种原始形态的基于群体智能的软件开发实践, 指出了基于群体智能的软件开发面临的主要问题和挑战, 并给出了对基于群体智能的软件开发理想形态的基本观点. 以这篇观点性论文为指导, 我们随后展开了一系列的研究工作, 具体包括“基于群体智能的拼图”(本文研究工作)、“基于群体智能的知识图谱构造”、“基于群体智能的程序缺陷修复”、“基于群体智能的程序融合”、“基于群体智能的机器人群体自组装”等方面, 并取得了一定的研究进展. “探索-融合-反馈”回路在这些研究工作中都展现出基础性和框架性的作用(由于篇幅限制, 细节信息不在本文展开). 因此, 我们对“探索-融合-反馈”回路对其他问题的适用性具有充分的信心.

## 6 总结与展望

本文提出了一种基于环境激发效应的群体协同式拼图问题求解方法. 该方法支持一组在地理上分布的人类玩家通过去中心化协同的方式求解拼图问题. 该方法的核心是一种包含探索、融合与反馈 3 个活动的回路. 该回路的核心输出制品是一个持续演化的群体观点图, 其将当前时刻玩家群体中每一个体的拼图结果实时地融合在一起, 形成一个结构化的信息制品; 群体观点图是反馈活动的基础. 我们开发了一个支持该方法的多人在线拼图环境, 并在其中进行了一系列受控实验, 以验证本文方法的可行性和有效性. 实验结果表明: (1) 在这一多人拼图环境中, 拼图求解时间大致正比于玩家数量的倒数, 且以玩家群体中的最强个体为基准点, 群体拼图效率能够实现 31.36%~64.57% 的提升; (2) 在该环境中, 最快完成拼图的玩家获得的反馈信息平均具有 86.34% 的准确率, 且随着群体规模的增加, 反馈信息在最快玩家拼图结果中的平均占比逐渐从 20% 增加到 45% 左右; (3) 相比于面对面协同的群体拼图求解方式, 该环境展示出更好的群体规模可扩展性, 且拼图结果总是具有 100% 的正确率, 而拼图问题自动求解算法平均只具有 52% 的正确率.

后续的研究工作将在两个方面展开. 在狭义上, 我们将继续研究基于群体智能的拼图问题求解方法. 我们计划招募更多的志愿者参与到更大规模的拼图问题求解实验中, 进一步验证本文方法的群体规模可扩展性与问题规模可扩展性; 并针对实验中发现的问题对本文方法进行持续的改进. 在广义上, 我们计划将从“基于群体智能的拼图”这一童话式场景中学习到的经验和知识应用到更具实际背景的复杂问题的群智求解上. 目前, 已经初步定位了两种复杂问题: 知识图谱构造问题和软件开发问题. 同时, 我们计划基于图的信息表示机制<sup>[46]</sup>, 开发一个一般性的信息融合与反馈平台, 以支持在该平台上快速开发面向特定问题的群智求解支撑环境.

**致谢** 感谢匿名审稿人和编委对论文初始提交版本提出的宝贵修改意见和建议.

## 参考文献

- 1 Torsvik T H. The Rodinia jigsaw puzzle. *Science*, 2003, 300: 1379-1381
- 2 Burkitt D P. Large-bowel cancer: an epidemiologic jigsaw puzzle. *J Natl Cancer Inst*, 1975, 54: 3-6
- 3 Brimacombe R. The structure of ribosomal RNA: a three-dimensional jigsaw puzzle. *Eur J Biochem*, 1995, 230: 365-383

- 4 Verweij M, Thompson M. Clumsy Solutions for a Complex World: Governance, Politics and Plural Perceptions. London: Palgrave Macmillan, 2006
- 5 Zhang W, Mei H. Software development based on Internet collective intelligence: feasibility, state-of-the-practice, and challenges. *Sin Chin Inform*, 2017, 47: 1601–1622 [张伟, 梅宏. 基于互联网群体智能的软件开发: 可行性、现状与挑战. *中国科学: 信息科学*, 2017, 47: 1601–1622]
- 6 Theraulaz G, Bonabeau E. A brief history of stigmergy. *Artif Life*, 1999, 5: 97–116
- 7 Karsai I. Decentralized control of construction behavior in paper wasps: an overview of the stigmergy approach. *Artif Life*, 1999, 5: 117–136
- 8 Susi T, Ziemke T. Social cognition, artefacts, and stigmergy: a comparative analysis of theoretical frameworks for the understanding of artefact-mediated collaborative activity. *Cogn Syst Res*, 2001, 2: 273–290
- 9 Dorigo M, Bonabeau E, Theraulaz G. Ant algorithms and stigmergy. *Future Gener Comput Syst*, 2000, 16: 851–871
- 10 Freeman H, Garder L. Apictorial jigsaw puzzles: the computer solution of a problem in pattern recognition. *IEEE Trans Electron Comput*, 1964, 13: 118–127
- 11 Berger R. The Undecidability of The Domino Problem. Washington: American Mathematical Society, 1966
- 12 Demaine E D, Demaine M L. Jigsaw puzzles, edge matching, and polyomino packing: connections and complexity. *Graphs Combin*, 2007, 23: 195–208
- 13 Goldberg D, Malon C, Bern M. A global approach to automatic solution of jigsaw puzzles. In: *Proceedings of the 18th Annual Symposium on Computational Geometry, Barcelona, 2002*. 82–87
- 14 Kong W, Kimia B B. On solving 2D and 3D puzzles using curve matching. In: *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai, 2001*. 583–590
- 15 Radack G M, Badler N I. Jigsaw puzzle matching using a boundary-centered polar encoding. *Comput Graph Image Process*, 1982, 19: 1–17
- 16 Wolfson H, Schonberg E, Kalvin A, et al. Solving jigsaw puzzles by computer. *Ann Oper Res*, 1988, 12: 51–64
- 17 Kosiba D A, Devaux P M, Balasubramanian S, et al. An automatic jigsaw puzzle solver. In: *Proceedings of the 12th IAPR International Conference on Pattern Recognition, Jerusalem, 1994*. 616–618
- 18 Makridis M, Papamarkos N. A new technique for solving a jigsaw puzzle. In: *Proceedings of International Conference on Image Processing, Atlanta, 2006*. 2001–2004
- 19 Nielsen T R, Drewsen P, Hansen K. Solving jigsaw puzzles using image features. *Pattern Recogn Lett*, 2008, 29: 1924–1933
- 20 Yao F H, Shao G F. A shape and image merging technique to solve jigsaw puzzles. *Pattern Recogn Lett*, 2003, 24: 1819–1835
- 21 Cho T S, Avidan S, Freeman W T. A probabilistic image jigsaw puzzle solver. In: *Proceedings of the 23rd IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, 2010*. 183–190
- 22 Pomeranz D, Shemesh M, Ben-Shahar O. A fully automated greedy square jigsaw puzzle solver. In: *Proceedings of the 24th IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, 2011*. 9–16
- 23 Yang X, Adluru N, Latecki L J. Particle filter with state permutations for solving image jigsaw puzzles. In: *Proceedings of the 24th IEEE Conference on Computer Vision and Pattern Recognition, Colorado Springs, 2011*. 2873–2880
- 24 Gallagher A C. Jigsaw puzzles with pieces of unknown orientation. In: *Proceedings of the 25th IEEE Conference on Computer Vision and Pattern Recognition, Providence, 2012*. 382–389
- 25 Sholomon D, David O E, Netanyahu N S. A genetic algorithm-based solver for very large jigsaw puzzles. In: *Proceedings of the 26th IEEE Conference on Computer Vision and Pattern Recognition, Portland, 2013*. 1767–1774
- 26 Sholomon D, David O E, Netanyahu N S. An automatic solver for very large jigsaw puzzles using genetic algorithms. *Genet Program Evol Mach*, 2016, 17: 291–313
- 27 Son K, Hays J, Cooper D B. Solving square jigsaw puzzles with loop constraints. In: *Proceedings of the 13th European Conference on Computer Vision, Zurich, 2014*. 32–46
- 28 Parunak H V D. A survey of environments and mechanisms for human-human stigmergy. In: *Proceedings of the 2nd International Workshop on Environments for Multi-Agent Systems, Utrecht, 2005*. 163–186
- 29 Heylighen F. Collective intelligence and its implementation on the web: algorithms to develop a collective mental map. *Comput Math Organ Theory*, 1999, 5: 253–280
- 30 Malone T, Laubacher R, Dellarocas C. The collective intelligence genome. *IEEE Eng Manag Rev*, 2010, 38: 38–52

- 31 Rosenberg L B. Artificial swarm intelligence, a human-in-the-loop approach to A.I.. In: Proceedings of the 30th AAAI Conference on Artificial Intelligence, Phoenix, 2016. 4381–4382
- 32 Rosenberg L B. Human swarms, a real-time method for collective intelligence. In: Proceedings of the 13th European Conference Artificial Life, York, 2015. 658–659
- 33 Lee J, Kladwang W, Lee M, et al. RNA design rules from a massive open laboratory. Proc Natl Acad Sci USA, 2014, 111: 2122–2127
- 34 Levy P. Collective Intelligence: Mankind's Emerging World in Cyberspace. Cambridge: Perseus Books, 1997
- 35 Woolley A W, Chabris C F, Pentland A, et al. Evidence for a collective intelligence factor in the performance of human groups. Science, 2010, 330: 686–688
- 36 Nielsen M. Reinventing Discovery: the New Era of Networked Science. Princeton: Princeton University Press, 2011
- 37 Maleszka M, Nguyen N T. Integration computing and collective intelligence. Expert Syst Appl, 2015, 42: 332–340
- 38 Brabham D C. Crowdsourcing as a model for problem solving: an introduction and cases. Convergence, 2008, 14: 75–90
- 39 Howe J. Crowdsourcing: a definition. 2006. [http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing\\_a.html](http://crowdsourcing.typepad.com/cs/2006/06/crowdsourcing_a.html)
- 40 Howe J. The rise of crowdsourcing. Wired, 2006, 14: 1–4
- 41 Doan A, Ramakrishnan R, Halevy A Y. Crowdsourcing systems on the world-wide web. Commun ACM, 2011, 54: 86–96
- 42 Estellés-Arolas E, González-Ladrón-de-Guevara F. Towards an integrated crowdsourcing definition. J Inf Sci, 2012, 38: 189–200
- 43 Kittur A, Chi E H, Suh B. Crowdsourcing user studies with Mechanical Turk. In: Proceedings of Conference on Human Factors in Computing Systems, Florence, 2008. 453–456
- 44 Valentine M A, Retelny D, To A, et al. Flash organizations: crowdsourcing complex work by structuring crowds as organizations. In: Proceedings of Conference on Human Factors in Computing Systems, Denver, 2017. 3523–3537
- 45 Kittur A, Nickerson J V, Bernstein M, et al. The future of crowd work. In: Proceedings of Conference on Computer Supported Cooperative Work, San Antonio, 2013. 1301–1318
- 46 Romero J C, Valdez M G. Using a graph based database to support collaborative interactive evolutionary systems. In: Recent Advances on Hybrid Approaches for Designing Intelligent Systems. Berlin: Springer, 2014. 581–591

## 附录 A 符号与定义说明

本文中使用的的基本符号和定义如下.

- $(a_i)_1^K$ : 数列  $a_1, a_2, \dots, a_K$ .
- $|S|$ : 集合或数列  $S$  中包含的元素的数量.
- 给定一个图  $G$ ,  $\mathcal{V}(G)$  表示其节点集合,  $\mathcal{E}(G)$  表示其边集合.
- 给定图  $G$  中的一个节点  $v \in \mathcal{V}(G)$ ,  $d(v)$  表示节点  $v$  的度, 即: 节点  $v$  参与到了几条边中.
- 给定图  $G$  中的一条边  $e \in \mathcal{E}(G)$ ,  $\mathcal{V}(e)$  表示由  $e$  两端的两个节点构成的集合.
- $\mathbf{1}(x)$ : 一个谓词函数. 当传入的谓词  $x$  为真, 该函数返回 1; 否则, 返回 0.

**定义A1** (有穷实数降序数列的  $\epsilon$ -最大差分上序列) 给定一个有穷实数数列  $(a_i)_1^K$ , 满足  $a_i \geq a_{i+1}$ ,  $i \in [1, K)$ . 给定实数常量  $\epsilon \geq 0$ . 称  $\lceil (a_i)_0^K \rceil^\epsilon$  为  $(a_i)_0^K$  的  $\epsilon$ -最大差分上序列, 当且仅当其满足如下条件:

- (1)  $\forall i \in [1, K) : (a_i - a_{i+1}) \leq \epsilon |a_i| \Rightarrow \lceil (a_i)_1^K \rceil^\epsilon = (a_i)_1^K$ ,
- (2)  $\exists i \in [1, K) : (a_i - a_{i+1}) > \epsilon |a_i| \Rightarrow \lceil (a_i)_1^K \rceil^\epsilon = (a_i)_1^J \wedge J \in [1, K) \wedge (\forall k \in [1, J) : (a_k - a_{k+1}) < (a_J - a_{J+1})) \wedge (\forall k \in [J+1, K) : (a_k - a_{k+1}) \leq (a_J - a_{J+1}))$ .

按照这个定义, 一个有穷实数降序序列会在一对特定的邻接点 (该对邻接点在序列中所有的邻接点中具有最大的差值) 上断裂, 形成两个序列; 除非所有临界点  $a_j$  和  $a_{j+1}$  的相对差值  $\frac{a_j - a_{j+1}}{a_j}$  都小于或等于常量  $\epsilon$ . 在后一种情况中, 该序列的  $\epsilon$ -最大差分上序列就是其自身; 否则, 该序列的  $\epsilon$ -最大差分上序列是断裂形成的前缀序列. 例如, 给定序列  $\langle 10, 9, 8, 7, 6, 3, 2, 1 \rangle$ , 对于任意  $\epsilon < 0.5$ , 该序列的  $\epsilon$ -最大差分上序列为  $\langle 10, 9, 8, 7, 6 \rangle$ ; 对于任意  $\epsilon \geq 0.5$ , 该序列的  $\epsilon$ -最大差分上序列为其自身. 又例如, 给定序列  $\langle 10, 9.9, 9.8, 9.7 \rangle$ , 对于任意  $\epsilon < \frac{1}{98}$ , 该序列的  $\epsilon$ -最大差分上序列为  $\langle 10 \rangle$ ; 对于任意  $\epsilon \geq \frac{1}{98}$ , 该序列的  $\epsilon$ -最大差分上序列为其自身.

**性质A1** 给定实数常量  $\epsilon \geq 0$ , 任何一个有穷实数降序数列具有唯一一个  $\epsilon$  最大差分上序列.

# Solving pictorial jigsaw puzzles via Internet-based collective intelligence

Bo SHEN<sup>1,2</sup>, Wei ZHANG<sup>1,2\*</sup>, Haiyan ZHAO<sup>1,2</sup>, Zhi JIN<sup>1,2\*</sup> & Yanhong WU<sup>3</sup>

1. *Key Laboratory of High Confidence Software Technology (Peking University), Ministry of Education of China, Beijing 100871, China;*

2. *Institute of Software, School of Electronics Engineering and Computer Science, Peking University, Beijing 100871, China;*

3. *School of Psychological and Cognitive Sciences, Peking University, Beijing 100871, China*

\* Corresponding author. E-mail: zhangw.sei@pku.edu.cn, zhijin@pku.edu.cn

**Abstract** Pictorial jigsaw (PJ) puzzles are a well-known leisure game for humans. Usually, a PJ puzzle game is played by one or several players face-to-face in physical space. In this paper, we focus on how to solve PJ puzzles in cyberspace by a group of physically-distributed players. We propose an approach to solve PJ puzzles using stigmergy-inspired Internet-based human collective intelligence. The core of the approach is a continuously-executing loop, namely, the EIF loop, which consists of three activities: exploration, integration, and feedback. During the exploration activity, each player tries to solve the PJ puzzle alone, without direct interaction with the other players. At any time, the result of a player's exploration represents a partial solution to the PJ puzzle and a set of rejected neighboring relationships between the puzzle pieces. The results of all players' explorations are integrated in real time through the integration activity, forming a continuously-updated collective opinion graph (COG). Through the feedback activity, each player is provided with personalized feedback information based on the current COG and the player's current exploration results, in order to accelerate his/her puzzle-solving process. Exploratory experiments show that: (1) Supported by this approach, the time to solve PJ puzzles is nearly linear with respect to the reciprocal of the number of players. Furthermore, compared with the best single players in the experiments, the puzzle-solving time decreases by 31.36%–64.57% on average for groups composed of 2 to 10 players. (2) Supported by this approach, the feedback information received by the best player in a group has an average precision of 86.34%, and, as the group size increases, the feedback information in the puzzle-solving result of the best player increases gradually from 20% to 45%. (3) This approach exhibits a better scalability with puzzle size than that of a face-to-face collaboration among ten players. Additionally, it always leads to 100%-accurate puzzle-solving results, whereas the results of the automated PJ-puzzle solver have an average accuracy of only 52%. It is envisaged that these results will provide useful information or opinions to facilitate a broader application of human collective intelligence in the Internet environment.

**Keywords** collective human intelligence, Internet, complex problem solving, pictorial jigsaw puzzle, information integration and feedback



**Bo SHEN** was born in 1994. He received his B.S. degree in Software Engineering from Northwestern Polytechnical University in 2016. Currently, he is a Ph.D. candidate in the School of Electronics Engineering and Computer Science at Peking University. His research interests include programming language, collaborative software development, and collective intelligence-based software engineering.



**Wei ZHANG** was born in 1978. He received his Ph.D. degree in Computer Science from Peking University in 2006. Currently, he is an associate professor in the School of Electronics Engineering and Computer Science at Peking University. His research interests include theories of collective intelligence, collective intelligence-based software engineering, and collective intelligence system design.



**Zhi JIN** was born in 1962. She received the Ph.D. degree in Computer Science from the National University of Defense Technology, in 1992. Currently, she is a full professor at Peking University. Her research interests include requirements engineering, knowledge engineering, and self-adaptive systems.



**Yanhong WU** was born in 1965. She received the Ph.D. degree in Psychology from Peking University, in 1997. Currently, she is a full professor in Psychology at Peking University. Her research interests include social and cultural neuroscience, memory inhibition, and attentional mechanisms.